DRES

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

SR 633

UNCLASSIFIED

Molecular Cloning and Sequencing of the 26S Region of Western Equine Encephalitis Virus Strain 71V-1658

By:

Fay L. Schmaltz, Les P. Nagata[1] and George Rayner
Medical Countermeasures Section

[1] Corresponding Author

January 1997

CRAD

Unlimited Distribution

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

SUFFIELD REPORT NO. 633

Molecular Cloning and Sequencing of the 26S Region of Western Equine Encephalitis Virus Strain 71V-1658

BY

Fay L. Schmaltz, Les P. Nagata[1] and George Rayner

Medical Countermeasures Section

Corresponding author[1]

ABSTRACT

A 3100 bp cDNA clone (pcDW-12) of the 26S region of western equine encephalitis (WEE) virus strain 71V-1658 was identified by dot blot hybridization from a cDNA library. The missing 5' end was PCR-amplified and engineered into pcDW-12 to obtained a full length clone of the 26S region. The resulting construct (XH-7) was restriction mapped and completely sequenced on both strands. Only eleven of the sixty-three nucleotide differences resulted in amino acid changes when the sequence was compared to WEE strain BFS-1703 (Hahn *et al*, 1988). In addition, the high degree of conservation of the structural proteins was maintained when compared to the N-terminal sequence of the E1 and E2 proteins of the McMillan strain of WEE (Bell *et al*, 1983). The conserved nature of the structural proteins would indicate that one strain of WEE should be able to cross-protect against all WEE strains. A 2.2 kb fragment at the 5' end of the genome was also cloned and sequenced, and demonstrated high homology to the available sequences for eastern equine encephalitis (EEE) and Highlands J (HJ) viruses, adding further evidence that the entire 5' nonstructural region of WEE was derived from EEE. In summary, the genetic cloning and sequencing of the WEE 26S region marks the critical first step in the generation of a subunit vaccine to WEE.

TABLE OF CONTENTS **Page No.**

Page No.

LIST OF TABLES AND FIGURES **Page No.**

INTRODUCTION

Western equine encephalitis virus is an alphavirus which is serologically related to Sindbis (SIN), HJ, Fort Morgan, Buggy Creek and Aura viruses. It appears to be a recombinant virus between SIN and EEE viruses (Hahn *et al*, 1988). WEE is endemic in western North America and subtypes/varieties have been isolated from Argentina (AG80-646), Brazil (BeAr 102091) and the former Soviet Union (Y62-33). In nature, WEE is transmitted to man and horses by mosquitoes (*Culex tarsalis* is the principal vector) from its amplifying host or reservoir of wild birds. Infection of man from the endemic cycle has resulted in only a limited number of infections in recent years. However, in the past, major epidemics of WEE have been recorded. The most extensive epidemic occurred in the western United States and Canada in 1941, with 3,336 recognized human cases and 300,000 cases of encephalitis in horses and mules (Peters and Dalrymple, 1990).

The incubation period for WEE is between 5-10 days. As described by Peters and Dalrymple, 1990, the infection varies from mild fever with headache to aseptic meningitis and encephalitis. The disease usually begins with the sudden onset of fever, malaise, headache, nausea, vomiting, and occasionally, respiratory symptoms. Vertigo, photophobia, lethargy, abdominal pain and myalgia are common symptoms. After a few days the headache becomes more intense, and in severe cases, disorientation, stupor, tremors, paralysis or coma may develop. Fatal cases usually succumb during the first week. Infants and children are more susceptible, and the onset may be more abrupt, often associated with seizures and convulsions. Pathological reflexes vary, but commonly include weakness and hypoflexia, and children often display nuchal rigidity, disturbed sensorium, involuntary movements and paralysis. After 10 days, patients usually begin a gradual convalescence. Most adults will recover fully, although infants and children are at higher risk of developing permanent damage (Peters and Dalrymple, 1990; Horne *et al*, 1991).

The pathogenicity of WEE resembles that of EEE, although it is less neuroinvasive and neurovirulent in both humans and laboratory animals. The primary pathological findings are located in the central nervous system and include multiple foci of necrosis. The case infectivity

ratio has been estimated to be around 1:1150 in adults and drops to 1:58 in children and 1:1 in infants. The case fatality rate is between 3-8%, although the infectious dose is not known (Peters and Dalrymple, 1990). Of note, out of seven reported cases of laboratory-associated/ acquired infections with WEE (associated with egg cultures, suckling mice and aerosols from lyophilized material), two deaths were recorded (Office of Biosafety, Laboratory Centre for Disease Control - LCDC, Canada). No direct human to human transmission of the virus has been recorded and the virus is not demonstrated in blood of man after the onset of disease (Peters and Dalrymple, 1990). Highlands J virus is antigenically related to WEE, and was previously referred to as the eastern variant of WEE. It does not cause clinical disease in humans or horses, although a case of encephalitis in a horse was attributed to HJ (Hayes and Wallis, 1977; Karabatsos *et al*, 1988).

Western equine encephalitis virus is an enveloped virus, containing a nucleocapsid (NC) which encases a single strand of positive-sense genomic RNA. The RNA is estimated to be near 12 kb in length, based on the size of the SIN virus genome (Strauss and Strauss, 1994). The nonstructural proteins are translated directly from the 5' two-thirds of the genomic RNA, while the structural proteins are translated from a 26S subgenomic mRNA encoded in the 3' end. The 26S region of WEE strain BFS 1703 has been cloned and sequenced (Hahn *et al*, 1988). The structural proteins (NC, E3, E2, 6K and E1) are processed by proteolytic cleavage from a single polyprotein. Apart from the NC, the other proteins are posttranslationally processed in the endoplasmic reticulum including glycosylation of the E1 and E2 proteins and the addition of lipid. Extrapolating from research with SIN and Venezuelan equine encephalitis (VEE) viruses, neutralizing domains were mapped primarily to the E2 protein, and to a lesser extent, to the E1 protein (reviewed in Roehrig, 1986).

An inactivated WEE vaccine is in Investigational New Drug (IND) status and requires a minimum of three injections as well as yearly boosters. Its major problem is the relatively poor and short-term immunogenicity to WEE that is produced. Currently, no antiviral to WEE has shown potential for treatment of infections, although a synthetic double stranded RNA, poly-ICLC, will be evaluated at DRES. In summary, a better vaccine needs to be developed for WEE. At the US Army Medical Research Institute Infectious Diseases(USAMRIID), research

is progressing towards the production of a live attenuated WEE vaccine via mutagenesis of an infectious clone of WEE virus. Live attenuated vaccines have proven to provide long-term immunity, and in many cases, protection is life-long. However, new types of problems can arise with these vaccines. The live attenuated VEE vaccine (TC83) has had trouble with reversion of the virus to a more virulent form (occurred in up to 20 % of the recipients of the vaccine - personal communication, J. Smith, USAMRIID). A second potential problem to be addressed is whether a live attenuated WEE virus will interfere with live attenuated VEE or EEE components of a trivalent vaccine. As an alternative, USAMRIID is looking at a baculovirus-produced subunit vaccine and microencapsulation for enhanced delivery of inactivated vaccine.

At DRES, research has focused on the development of subunit vaccines to WEE. Subunit vaccines involve inoculation of individuals with the protective component of the virus (envelope proteins). Subunit vaccines are generally safer than live vaccines, causing fewer complications during vaccination and no risk of reversion. They can be produced in a variety of expression systems (i.e. yeast, vaccinia, or as plasmid DNA), eliminating the need for high level biocontainment production facilities. Although multiple doses are usually required, recent advances in microencapsulation and DNA immunization have allowed the administration of a single dose which can induce a protective immune response. Liposomal-encapsulation, in conjunction with aerosol delivery, could prove effective in enhancing the immune response (Wong *et al*, 1992; Saravolac *et al*, 1996). The first step in the construction of a recombinant vaccine is to clone the relevant sequences from the viral genome. In this paper, we describe the steps in the cloning and DNA sequencing of the 26S region of WEE virus strain 71V-1658.

MATERIALS AND METHODS

CHEMICALS AND REAGENTS

Chemicals and reagents were purchased from Sigma Chemicals (St. Louis, MO), Aldrich (St. Louis, MO) and Fisher Scientific (Edmonton, AB). Restriction and DNA modification

enzymes were purchased from Gibco/BRL (Burlington, ON), Pharmacia (Baie d'Urfé, PQ) or New England Biolabs (Beverly, MA). Other suppliers include: GeneAmp™ polymerase chain reaction (PCR) reagents purchased from Perkin Elmer (Foster City, CA); Multiprime™ random primer labelling kits and radioisotopes from Amersham (Oakville, ON); avian myeloblastosis virus (AMV) reverse transcriptase from Life Sciences (Petersburg, FL); tissue culture media, sera, E-link™ kit and Photogene™ reagent from Gibco/BRL (Burlington, ON); and bacteriological media (Difco, Detroit, MI; Gibco/BRL, Burlington, ON). DNA primers were synthesized and gel purified at the Regional DNA Synthesis Laboratory (Calgary, AB) based on the published WEE 26S sequence (Hahn et al, 1988).

VIRUS CULTURE AND PURIFICATION

Tissue culture was maintained in accordance to established practices (Bird and Forrester, 1981). BHK-21 and Vero cells were obtained from American Type Culture Collections (ATCC - Rockville, MD) and USAMRIID. All manipulations of WEE were performed following Biocontainment Laboratory level 3 (BL-3) requirements at USAMRIID. A 10% suckling mouse brain (SMB) suspension of WEE strain 71V-1658 was kindly provided by Dr. Nick Karabatsos, Centers for Disease Control (CDC), Fort Collins, CO. A seed stock of 71V-1658 was made by inoculation of 4 roller bottles of Vero 19 cells (USAMRIID) with 70 µl total volume of the SMB suspension at a multiplicity of infection (MOI) of less than 0.1. Each roller bottle of cells (90-95% confluent) was incubated with the virus in 10 ml of Earles minimal essential media (EMEM) supplemented with 10 mM HEPES buffer, 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 50 µg/ml gentamycin containing 5% fetal calf serum (FCS) - 5% EMEM. Incubation was continued for 1 hr at 37°C and an additional 40 ml of 5% EMEM was added to each roller bottle. The cytopathic effect (CPE) was allowed to proceed to 80-85% (approximately 3 days). The lysate was harvested, debris removed by centrifugation at 10,000 x g, and the supernatant aliquotted and stored at -70°C. Virus stocks were prepared by infecting 20 roller bottles with the WEE seed stock at a MOI of 10. After cell lysis (80-85% CPE), the virus was precipitated from cleared supernatant by the addition of polyethylene glycol (PEG) MW 6000 to 7%(w/v) and NaCl to 2.3%(w/v). The virus suspension was gently stirred for 15-20 hr at 4°C, followed by centrifugation at 10,000 x g for 30 min at room

temperature (RT). The pellets were resuspended in a total volume of 10 ml of phosphate-buffered saline (PBS - 0.01 M sodium phosphate buffer, 27 mM KCl, 0.137 M NaCl, pH7.4), and 5 ml of virus suspension was carefully layered onto 30 ml of a 20-60%(w/w) continuous sucrose gradient in a SW 28 tube (Beckman, Mississauga, ON). The tube was centrifuged at 100,000 x g (27,500 RPM in SW 28 rotor) for 3.5 hr at 4°C. The virus was harvested by removal of the virus band with a pasteur pipette. The virus suspension was then diluted at least 1:3 with Hanks Buffered Saline Solution (HBSS - Bird and Forrester, 1981), and centrifuged at 100,000 x g for 1 hr at 4°C. The virus pellet was resuspended in HBSS or PBS (5-10 ml), aliquotted and stored at -70°C.

ISOLATION OF RNA

Viral RNA used in library constructions to HJ strain B320 and WEE strain 71V-1658 were a kind gift of Dr Dennis Trent, CDC (Fort Collins, CO). Subsequent batches of RNA were prepared from WEE virus strain 71V-1658 cultured on Vero or BHK-21 cells. The virus was lysed by incubation in 0.5 % sodium dodecyl sulfate (SDS) and RNA extracted using the cesium chloride (CsCl)/guanidium isothiocyanate method as described (Sambrook *et al*, 1989), with the RNA being pelleted through a 5.7 M CsCl cushion. The RNA was aliquotted, precipitated by the addition of 1/10 volume of 3 M sodium acetate and 2 volumes of ethanol (International Biotechnologies Inc. - IBI, New Haven, CT), then stored at -70°C. Prior to use, the RNA was pelleted 15 min at 11,600 xg in a microfuge at 4°C, washed with 80%(v/v) ethanol and dried in a Savant Speed-Vac AES 1000 (Savant Instruments, Hicksville, NY). The pellet was dissolved in nuclease-free water (Promega, Madison, WI; Gibco/BRL, Burlington, ON), and alliquots were stored at -70°C. Integrity of the RNA was checked on formaldehyde agarose gels as described by Sambrook *et al*, 1989. Aerosol-free, RNase-free pipette tips (Gordon Technologies, Mississauga, ON) siliconized RNase-free microfuge tubes, use of diethylpyrocarbonate (DEPC)-treated dH_2O for buffers and aseptic techniques were used for all manipulations and reactions. Gel boats were cleaned with 70%(v/v) ethanol and DEPC-treated water immediately before use.

cDNA LIBRARY CONSTRUCTION

Libraries to HJ strain B320 or WEE strain 71V-1658 were made by Invitrogen (San Diego, CA) using the following protocol: an oligo-dT primer was used for reverse transcription of 5 μg of viral RNA to cDNA. After synthesis of the second DNA strand, linkers were added and the dsDNA was size selected to greater than 2,000 bp. The sized, electroeluted DNA was ligated into the *BstXI* site of prepared pcDNAII vector, and electroporated into electrocompetent DH1αF' *Escherichia coli* cells. The cells were amplified by growing on LB agar containing 50-100 μg/ml ampicillin.

MANIPULATION OF DNA

Well publicized methods were used for handling and manipulation of DNA (Sambrook *et al*, 1989; Ausubel *et al, 1995).* These include the use of aseptic technique for the handling bacteria and tissue culture. Where possible, high grade reagents, nuclease-free water (Promega, Madison, WI), siliconized microfuge tubes, sterile solutions and disposables were used. Phenol/chloroform extractions used an equal volume of a mixture of water-saturated distilled phenol (Gibco/BRL, Burlington, ON), chloroform and isoamyl alcohol in a ratio of 25:24:1 ratio. This was followed by ethanol precipitation using 1/10 to 1/20 volume sodium acetate (3 M) and 2-2.5 volumes molecular biology grade ethanol (IBI, New Haven, CT). The DNA could be stored indefinitely at -20°C or pelleted 10 min at 11,600 x g, washed once with 80%(v/v) ethanol and dried in a Speed-Vac™ AES 1000 centrifugal vacuum drier (Savant Instruments, Hicksville, NY). Nucleic acid was quantitated by measurement of the absorbance of an aliquot of the sample at 260 nm using a Unicam 8730 spectrophotometer (Mississauga, ON).

RESTRICTION ANALYSIS OF DNA

Restriction digests were performed by dilution of DNA (up to 5 μg) in dH_2O and with the addition of 1 μl RNase (10 mg/ml, heat-inactivated by boiling 20 min), 2 μl restriction endonuclease buffer (10x) in a total volume of 20 μl. Up to 2 μl of restriction endonuclease(s) were added (1-5 U/μg DNA) and the reaction allowed to incubate at the specified temperature

(usually 37°C). For restriction mapping, digestions were allowed to proceed for 2-3 hrs, while for DNA fragment isolation, the digests were incubated for exactly 1 hr. Three µl of 10x gel loading dye [50%(v/v) glycerol, 25%(w/v) ficoll, 0.25% (w/v) bromphenol blue, 0.25%(w/v) xylene cyanol FF] was added, and the mixture loaded onto an agarose gel for electrophoresis. The gels were cast in removable plastic trays (11 X 14 cm) to a thickness of 3-5 mm, using 100 ml of a 1.0-1.5%(w/v) agarose (Gibco/BRL, Burlington, ON; Seakem™ or Seaplaque™ - FMC Bioproducts, Rockland, ME) solution made up in gel running buffer. The gel buffers used were lx Tris borate EDTA (TBE) (90 mM Tris-base, 90 mM boric acid and 2 mM EDTA) or 1x Tris-Acetate EDTA (TAE) (40 mM Tris-base, 20 mM sodium acetate, 0.2 mM EDTA, adjusted to pH 7.4 with glacial acetic acid). The agarose was dissolved by microwave heating of the agarose/buffer suspension (3-4 min on 50% power), and the solution cooled to 40-50°C before 5 µl of 10 mg/ml ethidium bromide (EtBr) was added. The TBE gels were generally run at 80-120 V and the TAE gels run at 60-80 V in the appropriate 1x gel buffer (Sambrook *et al*, 1989; Ausubel *et al, 1995)*.

LABELLING OF GENE PROBES TO HJ AND WEE

For the screening of the HJ library, the HJ-3' and WEEPRO oligonucleotides were labelled using the E-link kit (Gibco/BRL, Burlington, ON) following the manufacturer's instructions. The oligonucleotide probes were also 5' end-labelled using γ-[^{32}P]ATP (Amersham, Oakville, ON) and polynucleotide kinase as described in Sambrook *et al* (1989). For WEE library screening, DNA fragments were generated using standard PCR reaction conditions (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.0 mM $MgCl_2$, 200 µM of each dNTP, 2.5 U Amplitaq™, 1.0 µM of each primer) as described by Perkin Elmer (Foster City, CA). Optionally, 1 U of Perfect Match™ (Stratagene, La Jolla, CA) could be added to the reaction mixture. Fifty µl reactions were overlaid with mineral oil or with PCR Gems™ in 500 µl microfuge tubes. The reaction program consisted of 25 cycles of 1 min at 94°C, 2 min at 37°C and 3 min at 72°C, followed by a final 7 min extension at 72°C in a Perkin Elmer DNA thermocycler (Foster City, CA). The primer pairs used to generate PCR fragments in the 26S region of WEE were WEEPRO/WEEP0, WEEP5/WEEN5, WEEP7/WEEN7 and WEE3'/WEEN0 (see Table 1). The reaction mix was loaded onto a 1%(w/v) TAE agarose gel containing 0.5 µg/ml EtBr.

DEAE 81 paper (Whatman, Hillsboro, OR) was used to trap the band of interest by placing an incision ahead of the DNA band of interest. A strip of DEAE 81 paper was inserted into the slot and electrophoresis continued for 15-30 min. The DEAE 81 paper containing the trapped DNA was then washed with low-salt wash buffer (10 mM Tris-HCl, pH 8.0, 0.10 M NaCl, 1 mM EDTA), followed by elution of the DNA with 50-100 µl of elution buffer (10 mM Tris-HCl, pH 8.0, 1.5 M NaCl, 1 mM EDTA). The eluant was phenol/$CHCl_3$ extracted, followed by ethanol precipitation (Danner, 1982; Sambrook *et al*, 1989). Alternatively, the fragment was isolated from low melting agarose gels by excision of the fragment, and subsequent phenol extraction of the melted agarose (Sambrook *et al*, 1989; Ausubel *et al,* 1995).

After washing and drying, the PCR fragment was labelled using the Multiprime™ random primer labelling kit (Amersham, Oakville, ON). Briefly, 25 ng of DNA in 10 µl of dH_2O was boiled 2 min and quickly cooled on ice. The following items were added: 11 µl dH_2O; 4 µl of each of dATP, dGTP and TTP; 5 µl buffer concentrate; 5 µl primer/BSA; 5 µl $\alpha[^{32}P]$-dCTP in tricene (3,000 Ci/mmole, 50 µCi - Amersham, Oakville, ON); and 2 µl of DNA polymerase I Klenow fragment (1U/µl). The reaction was incubated 30-60 min at 37°C, after which 50 µl of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) was added. The free nucleotides removed by passing the sample through a 1.0 ml G-50 Sephadex™ (Pharmacia, Baie d'Urfé, PQ) spun-column as described (Sambrook *et al*, 1989) or through the use of Nuctrap™ Push Columns (Stratagene, La Jolla, CA) following the manufacturer's instructions.

PREPARATION OF COLONY AND DOT BLOTS

Colony hybridization was based on the original procedure of Grunstein and Hogness (1975) as described in Sambrook *et al* (1989) to liberate the DNA from bacterial colonies and bind it to nitrocellulose filters (Hybond-C Extra - Amersham, Oakville, ON). Pieces of Whatman 3MM paper were cut to an appropriate size and shape to fit into the bottoms of three glass trays. The filter paper was saturated with one of the following solutions: 10% SDS; denaturing solution (0.5 N NaOH, 1.5 M NaCl); or neutralizing solution (1.5 M NaCl, 0.5 M Tris Cl, pH 7.4). Bacteria to be screened were plated out onto appropriate media (LB agar with 50-100 µg/ml ampicillin) and grown overnight (O/N). The nitrocellulose filters were placed on top of the 1-

2 mm colonies in the 90 or 150 mm plates and left 30-60 sec. The position of the filter relative to the plate was indicated by using a 18 or 20 g needle to puncture holes through the filter and into the plate. These holes could then be used to re-align the plate to an autoradiograph of the filter after hybridization, allowing positive colonies to be selected from the plate. The filters were lifted carefully from the plate and placed colony side up on the SDS impregnated 3MM paper. After the first filter had been exposed to the SDS solution for 5 minutes, it was transfered to the second sheet of 3MM paper, saturated with denaturing solution for 5 minutes. The filters were then transferred to the third sheet of 3MM paper, saturated with neutralizing solution, and left for 5 minutes. They were then placed on a clean sheet of Whatman 3MM paper (colony side up) and the surface of the filter was blotted once with a second sheet of 3 MM paper. The filters were transferred to a tray containing 2x SSC (0.3 M sodium chloride and 0.03 M sodium citrate, pH 7.0) and gently agitated for 5 min. They were dried at RT for at least 30 minutes, sandwiched between two sheets of 3MM paper and clamped between glass plates. They were baked for 1-2 hours at 80°C in a vacuum oven. Any filters not used immediately in hybridization reactions were sealed in plastic and stored at -70°C.

Dot blots were prepared for more accurate screening of recombinant plasmids with labelled probes (Kafatos *et al,* 1979). Nylon membranes (Hybond-N^{TM} - Amersham, Oakville, ON) were spotted with 1-2 µl of plasmid preparations along with appropriate controls. After air drying, the membranes were covered with plastic wrap and placed face down on a Fotodyne Model 3-3500 UV transilluminator (New Berlin, WI) for 5 min on the long UV setting. The membranes were denatured similarly as for colony filters, using 3 min in denaturing solution, 3 min in neutralizing solution and equilibrated for 5 min in 2x SSC (Grunstein and Hogness, 1975; Sambrook *et al*, 1989).

HYBRIDIZATION TO LABELLED DNA PROBES

Hybridization and washing with the HJ3' and WEEPRO E-link oligonucleotide probes followed the manufacturer's directions. For [^{32}P]-labelled probe screening, nitrocellulose filters and nylon membranes were hybridized utilizing formamide solutions (Denhardt, 1966, McConaughy *et al,* 1969). The membranes were prehybridized for 1 hr at 42°C in sealed

plastic bags containing 10-20 ml Hood buffer [50%(v/v) formamide (deionized); 20 mM sodium phosphate buffer (pH 6.7); 5x SSC (0.75 M sodium chloride, 0.075 M sodium citrate, pH 7.0); 7%(w/v) SDS; 1 %(w/v) PEG MW 20,000; and 0.5%(w/v) nonfat powdered milk (Carnation)]. For large numbers of filters, teflon mesh was used to separate the layers of filters. Hybridization was carried out in 5-10 ml Hood buffer containing 10^7 cpm of radiolabelled probe (Sambrook *et al*, 1989). After hybridization, filters or membranes were transferred to polyethylene trays and washed with two quick changes of 2x SSC, 1%(w/v) SDS. They were then washed with 2 changes of 2x SSC, 1%(w/v) SDS at 42-60°C, 20 min each. A final wash of 20 min at RT in 0.1x SSC, 1%(w/v) SDS with gentle agitation was performed, and the filters or membranes were sealed in plastic for exposure with Kodak XAR-5 X-ray film (Rochester, NY).

PREPARATION OF PLASMID DNA

A scaled-up version of the Holmes and Quigley (1981) procedure for rapid preparation of bacterial plasmids was followed. Clones identified by colony hybridization were cultured O/N in 2-5 ml of LB broth (10 g tryptone, 5 g yeast extract, 5 g NaCl in 1.0 liter dH_2O) with 50-100 µg/ml ampicillin. The cells were pelleted for 30 min at 3,000 RPM (Beckman TJ-6 bench top centrifuge - Mississauga, ON) and the supernatant discarded. The cell pellet was resuspended in 200 µl of STET buffer [8%(w/v) sucrose, 5% triton X-100, 50 mM EDTA and 50 mM Tris-HCl, pH 8.0] and the suspension transferred to a 1.5 ml microfuge tube. Twenty-five µl of freshly prepared lysozyme (10 mg/ml) was added and the samples boiled 1 min. The tubes were then spun for 15 min at 11,600 x g in a microfuge at RT. The supernatant was transferred to a clean 1.5 ml microfuge tube, phenol-chloroform extracted, and precipitated with an equal volume of isopropanol (Sambrook *et al*, 1989). The tube was placed at -20°C for a minimum of 30 min, and then pelleted at 11,600 x g. The pellet was washed, dried and dissolved in 100-200 µl sterile dH_2O.

Large batches of *E. coli* plasmid were isolated using the Qiagen kit (Chatsworth, CA), a method based on the alkali lysis method of Birnboim and Doly (1979), following the manufacturer's instructions. A second method used for large-scale plasmid preparation was by a procedure

involving lysis with triton X-100 (Davis et al, 1980). Cells were grown O/N in LB broth containing 50-100 μg/ml ampicillin (0.5-1.0 liter), then pelleted for 10 min at 10,000 x g (Beckman JA-14 rotor - Mississauga, Ont.). The pellet was washed in TES (50 mM Tris-HCl, pH8.0, 50 mM NaCl and 5 mM EDTA) and then resuspended in 1.5 ml of cold sucrose buffer (25% sucrose, 50 mM Tris-HCl, pH8.0 and 1 mM EDTA). After 10 min on ice, 0.15 ml of 10 mg/ml lysozyme (fresh, in dH2O) was added and the cells were left for an additional 10 min on ice. Next, 0.6 ml of 0.25 M EDTA was added, and the cells were incubated for a further 10 min on ice. Slowly, 5.25 ml of triton lysis buffer [2%(v/v) triton X-100, 50 mM Tris-HCl, pH 8.0 and 60 mM EDTA] was added with gentle stirring, and the mixture left on ice for 15-20 min. Cell debris was removed by centrifugation for 25 min at 37,000 x g (Beckman J2-21, JA20 rotor - Mississauga, ON) after which the volume of the supernatant was adjusted to 9 ml with TE buffer. Nine grams of molecular biology grade CsCl (Gibco/BRL, Burlington, ON) was dissolved in each sample followed by the addition of 0.1 ml EtBr (10 mg/ml in dimethyl sulfoxide). The solution was transferred to Beckman heat-seal tubes, filled with paraffin oil, sealed and centrifuged for 18-24 hrs at 247,000 x g (Beckman L8-70, 70.1 Ti rotor - Mississauga, ON). The plasmid band was collected by puncture of the tube using a 18 g needle and 3 ml syringe was used to withdraw the band. The DNA was extracted with dH2O-saturated N-butanol (Sambrook *et al,* 1989) and dialyzed against TE buffer.

DNA SEQUENCE ANALYSIS

Manual sequencing was performed using the Sequenase™ ver 2.0 reaction kit (Amersham, Oakville, ON). The plasmid template (i.e. pcDW-6, pcDW-12, XH-7) was previously purified using one of the large-scale plasmid purification methods. A 1 μg/μl plasmid stock in dH_2O was required. Five μg of plasmid template was denatured by the addition of 1/10 volume 2 N NaOH and incubated 30 min at 37°C. This was followed by neutralization with 1/10 volume of 3 M sodium acetate. The sample was then precipitated by the addition of 4 volumes of 95%(v/v) ethanol, left O/N at -20°C, pelleted, washed and resuspended in dH_2O. For each sequencing reaction, 5 μg of denatured template in 7 μl dH_2O was mixed with 1 μl of primer (0.5 pmol) and 2 μl of Sequenase™ reaction buffer. The mixture was incubated for 2 min at 65°C, then slow cooled in a beaker of water to around 30°C. One μl of 1.0 M dithiothreitol

(DTT), 2 µl labelling mix (diluted 1/5) and 0.5 µl α[^{35}S]dATP in tricene buffer, >1000 Ci/mmol (10 µCi/µl - Amersham, Oakville, ON) were added to the primer/template mixture. Aliquots of the primer/template mixture were added (3.5 µl) to each of four tubes containing 2.5 µl of the appropriate dideoxy-terminator mix (ddGTP, ddATP, ddTTP or ddCTP) which had been previously prewarmed at 37°C. The reactions were incubated 3-5 min at 37°C after which 4 µl of dye-stop solution was added to each tube. Reaction tubes were heated 2 min at 80°C immediately prior to loading on the gel. After the wells were flushed, 2-3 µl of each reaction mix was loaded in sets of four dideoxy-terminators (GATC) using flat Gel-Loader pipette tips (Gordon Technologies, Mississauga, ON; National Scientific, Claremont, CA) and a Gilson Pipetteman^{TM} P-20 pipettor (Mandel Scientific, Guelph, ON). After a 2-4 hr run time, a second set of the samples could be loaded into any remaining wells. Forty well sharktooth combs were used for 40 cm gels. A number of different gel formats were run, but generally consisted of 8% acrylamide (1:20 bis-acrylamide:acrylamide ratio), 8 M urea and 1x TBE buffer. Forty cm and 90 cm, 0.4 mm thick gels using Gibco/BRL (Burlington, ON) S0 and S2 sequencing units, respectively, were run at 40-50 W constant power on BioRad (Mississauga, ON) 3000 power supplies. When the bromphenol blue dye was near the bottom, the gel was removed, fixed with three changes of 10%(v/v) methanol/10%(v/v) acetic acid (5 min per change). The gel was rinsed with several changes of dH_2O, followed by transfer to a sheet of Whatman 2 or 3 MM filter paper. Ninety cm gels were cut into two equal pieces prior to transfer to the Whatman filter paper. The gel was covered with plastic wrap, and dried on a gel drier under vacuum for 30 min at 80°C. The plastic wrap was removed, and the dried gel and was exposed to Kodak (Rochester, NY) XAR-5 X-ray film (14" x 17") with a Dupont Cronex^{TM} intensifying screen (Mississauga, ON). The X-ray was developed in Kodak (Rochester, NY) GBX developer and fixer and analysed on a Speed Reader^{TM} semi-automated DNA sequence imager (Intelligenetics, Mountain View, CA). The edited sequence files were transferred to text format using PC/Gene^{TM} (Intelligenetics, Mountain View, CA) DNA analysis software, prior to transfer to the Lasergene^{TM} biocomputing software (DNASTAR Inc., Madison, WI) installed on a Power Macintosh^{TM} 8100/80 (Apple Computer, Cupertino, CA) for sequence assembly.

Automated sequencing was performed using an Applied Biosystem Instruments (ABI) 373 automated fluorescent sequencer and the ABI Prism™ Dye Terminator Cycle Sequencing kit (Perkin Elmer, Foster City, CA). Both Amplitaq™ and Amplitaq FS™ were used on Qiagen purified dsDNA plasmid templates according to the manufacturer's instructions. The reactions were run on a Perkin Elmer thermocycler (Foster City, CA) using 30 cycles of: 96°C for 30 sec; 45°C for 30 sec and 60°C for 4 min. Reactions were purified on Centri-Sep™ columns (Princeton Separations, Adelphia, NJ), dried and run on an ABI 373 automated sequencer (USAMRIID, Regional DNA Synthesis Laboratory). Chromatographic data obtained was downloaded onto the MacIntosh 8100/80 and edited in SeqEd™ ver 1.0.3 (ABI, Foster City, CA), or transferred directly to Lasergene™ (DNAStar, Madison, WI) for sequence assembly. Thermosequense™ reactions were also compared to the Amplitaq FS™ according to the manufacturer's protocol (Amersham, Oakville, ON), however, the results were no better than the Amplitaq FS™ (Perkin Elmer, Foster City, CA).

CONSTRUCTION OF THE FULL-LENGTH 26S REGION OF WEE

A clone, pcDW-12, containing a 3100 bp fragment of the 26S region of WEE strain 71V-1658 was identified from the WEE library by hybridization to WEE probes, restriction mapping and DNA sequencing. The 5' end of the 26S region was synthesized from the WEE 71V-1658 RNA using standard PCR reaction conditions, and the WEE5'Sst and WEEP3 primers as outlined in Fig. 1. Before the PCR reaction was run, a reverse transcription (RT) of the RNA template was completed in the same tube using the PCR reaction conditions by programming in a 42°C "soak" cycle before the start of the thermocycling reaction. The changes incorporated were as follows: the primers were allowed to slow anneal to the RNA template, followed by the addition of 18 U of AMV reverse transcriptase (Life Sciences, Petersburg, FL) and 2.5 U Amplitaq™ polymerase (Perkin Elmer, Foster City, CA). The reaction was incubated for 30-60 min at 42°C, and then the standard PCR cycling was programmed to run. The PCR products were then analyzed on agarose gels. An 1800 bp fragment of the estimated size (Hahn *et al*, 1988) was isolated as previously described. The fragment was then restriction digested with *NcoI* and *SstI*, gel purified, and ligated into the *SstI/NcoI* sites of plasmid, phT3T7BM+ (Boehringer Mannheim, Laval, PQ). To expand on the details, the *SstI/NcoI* digested

phT3T7BM+ (gel purified, 50-100 ng) was added to the *SstI/NcoI* 1800 bp PCR fragment generated from WEE5'Sst/WEEP3 primers (1:3, 1:1 and 3:1 molar ratios) in a total volume of 15 µl dH_2O. Controls with digested plasmid DNA only and insert DNA only were also included. The DNA was slow annealed from 65°C, followed by the addition of 4 µl of 5x ligase buffer (Gibco/BRL, Burlington, ON) (250 mM Tris-HCl, pH 7.6, 50 mM $MgCl_2$, 25%w/v PEG 8000, 5 mM ATP, 5 mM DTT) and 1 U T4 DNA ligase (Gibco/BRL, Burlington, ON; Pharmacia, Baie d'Urfé, PQ). The reaction was incubated for 4 hr at 15°C in a refrigerated waterbath. Fifty µl of TE buffer was added and 2-5 µl of diluted ligation mixture was used to transform competent *E. coli* cells.

The transformation protocol of Hanahan (1983) was followed. Competent DH5α (maximum efficiency) *E. coli* cells were used, and purchased from Gibco/BRL (Burlington, ON). Briefly, 2-5 µl of diluted ligation reaction was added to 100 µl of cells in pre-chilled 15 ml snap cap tubes. The tubes were left for 30 on ice, heat shocked for 45 sec at 42°C and then cooled 2-3 min on ice. One ml of 2x YT [1 %(w/v) tryptone, 1 %(w/v) yeast extract, 0.5 %(w/v) NaCl] or SOB [2 %(w/v) tryptone, 0.5 %(w/v) yeast extract, 8.2 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$] broth was added and the tubes incubated at 37°C for 30 min with gentle agitation. The cell suspension (100 µl) was plated out onto a LB plate containing 50-100 µg/ml ampicillin.

A recombinant plasmid containing the 1800 bp WEE 5' end (phTW5'-3) was identified by restriction analysis of rapid plasmid preparations and confirmed by fine restriction mapping and DNA sequence analysis. The *XbaI/NcoI* fragment of phTW5'-3 (WEE 5' 1800 bp recombinant in phT3T7BM+) was ligated into the *XbaI/NcoI* sites of pcDW-12. After transformation into DH5α, a full-length recombinant containing the complete WEE 71V-1658 26S region was identified (XH-7) for further analysis. Additional cDNA constructs to WEE strain 71V-1658 were made using primers WEE5'Sst and WEE3'HindIII (see Table 1) for the RT-PCR amplification and directional cloning of the full-length WEE 26S into Invitrogen vector pRSETA (cloned into the *SstI* and *HindIII* sites).

PCR AMPLIFICATION OF THE NONSTRUCTURAL REGION OF WEE

A PCR cloning strategy was employed to produce a set of three overlapping fragments (WNST-A, B and C) for the nonstructural region of WEE71V-1658 (Fig. 2). RT-PCR reactions were carried out using three sets of PCR primers as listed in Fig. 2. For WNS-A fragment, two different combinations of primers were made, using a SIN5' or EHJ5' upstream primer, respectively. RT-PCR amplified cDNA was gel purified, and ligated into the modified EcoRI site of the pCRII vector supplied with the TA-cloning system™ (Invitrogen, San Diego, CA). The recombinant plasmid was transformed into *E. coli InvαF'* competent cells as described (Invitrogen, San Diego, CA). Isolates were screened by restriction mapping rapid plasmid preparations and by DNA sequencing using the T7 and M13 reverse primers (Invitrogen, San Diego, CA).

RESULTS

HJ AND WEE LIBRARY SCREENING

Plaque assays of the batches of WEE 71V-1658 grown at USAMRIID were completed (data not shown). The titre of the 10% SMB of WEE 71V-1658 was 6.1 X 10^8 pfu/ml. The seed stock was 4.1 X 10^7 pfu/ml. The quality of cDNA libraries is dependent foremost, on the integrity of the RNA starting material. RNA gels of HJ virus strain B320 and the initial batch of WEE 71V-1658 (CDC) indicated some degradation in the samples had occurred (data not shown). This is reflected in the large numbers of small inserts (<1000 bp) encountered when screening isolates of HJ library for insert size (see Fig. 3 and Table 2). This was surprising since Invitrogen indicated that the cDNA synthesis had attained an excellent size distribution, ranging from 100 bp to >8,000 bp. In addition, the cDNA was size selected (>2,000 bp) prior to ligation and transformation into the pcDNA II vector. In any event, we used oligonucleotide probes to the HJ 3' end (Ou *et al*,1981) and the promoter region (WEEPRO primer) of WEE (see Table 1). As genomic sequence information for HJ was limited to the 3' and 5' ends, we had speculated that if HJ and WEE were related, the 26S promoter area could show conservation of the genome sequence and cross-reactivity to HJ. A new, nonradioactive

procedure (E-link) was used to label HJ3' and WEEPRO oligonucleotide probes to screen isolates of interest (26S region). However, better hybridization results were acheived using probes conventionally labelled at the 5' end of the oligonucleotide with γ-[^{32}P]ATP and polynucleotide kinase. Nevertheless, over 100 isolates were screened by restriction mapping (see Fig. 3) and oligonucleotide probe reactivity (see Table 2). The size distribution of the inserts was as follows: 0-500 bp - 32%; 500-1000 bp -35%; 1000-2000 bp - 26%; and >2000 bp - 7%. Non-reactivity with the probes would only indicate that the isolate did not carry an insert that extended into the limited target region of the probe, and could not be interpreted as the isolate carrying no HJ DNA. A number of HJ clones were identified for DNA sequence analysis including: A19, A24, A35, B23, C12, D1 and D2). Further work was not completed with Highlands J isolates, as the emphasis shifted to the WEE strain 71V-1658 library.

As previously mentioned, the initial batch of WEE 71V-1658 showed degradation by RNA formaldehyde gel electrophoresis (data not shown), however, subsequent batches of genomic RNA of WEE 71V-1658 were very good (aliquots stored in ethanol at -70^{o}C). Sequence data was available to the entire 26S region of WEE strain BFS 1703 (Hahn *et al,* 1988) and was used to design a set of primers that could be used for PCR amplification, as well as DNA sequence analysis of WEE strain 71V-1658 (see Table 1). The fragments generated by PCR were designed to overlap at unique restriction sites, and allowed PCR generated fragments to be substituted into the WEE 26S clones. This was used to our advantage in the subsequent genetic engineering of a full-length 26S clone, as a complete clone could not be isolated from the Invitrogen library. Rapid plasmid isolates were screened by dot blot analysis using WEEPRO/WEEPO, WEEN5/WEEP5, WEEN7/WEEP7 and WEENO/WEE3' PCR generated fragments as probes (see Fig. 4). The results from dot hybridization are summarized in Table 3. Dot blot hybridization was used to quickly size inserts in the WEE library. The PCR probes selected were spaced along the 26S gene segment. We theorized that a full-length 26S insert would hybridize with all four PCR probes, while smaller the inserts would only hybridize to one or two probes. No isolates were identified which hybridized to all four probes, but a number of the largest isolates hybridized to three of the four probes. These included pcDW-1, 4, 6, 7 and 12 (data not shown). The size distribution of inserts in the pcDNA II vector was similar to that obtained with the HJ library, with not many large inserts obtained (>2,000 bp).

The largest clones identified were pcDW-6 and pcDW-12, which were estimated to be 2900 and 3100 bp, respectively, by restriction analysis on agarose gels (see Table 3). This was still 25-30 % shorter then the estimated 4100-4200 size of the 26S region of WEE. DNA sequence analysis confirmed that both clones pcDW-6 and pcDW-12 were truncated from the 5' end of the WEE 26S region at nucleotide (nt) 1798 and nt 1446 (BFS 1703 numbering system - Hahn *et al*, 1988), respectively. However, since the WEE sequencing primers were designed to produce fragments which terminated at unique restriction sites (BFS 1703 WEE 26S sequence - Hahn *et al*, 1988), a strategy was developed to place a PCR synthesized 5' fragment onto the end of pcDW-12 as outlined in Fig. 1.

CONSTRUCTION OF THE FULL-LENGTH 26S REGION OF WEE

RT-PCR was used to synthesize the WEE 26S 5' end using the WEE5'Sst and the WEEP3 primers. A fragment of the expected size (1800 bp) was isolated from agarose gels and ligated into the *Sst1* and *Nco1* sites of phT3T7BM+ (data not shown). A number of recombinant plasmids containing the 1800 bp WEE 5' end were screened by restriction analysis of rapid plasmid preparations, and confirmed by fine restriction mapping and DNA sequence analysis (phTW5'-3, 4 and 6 - data not shown) Cloning into phT3T7BM+ allowed us to excise the same 1800 bp fragment with the addition of a *Xba1* site at the 5' end. The *XbaI/NcoI* fragment of phTW5'-3 was ligated into the *XbaI/NcoI* sites of pcDW-12 and full-length 26S region clones (XH-7, 12) were screened for by restriction analysis of rapid plasmid preparations (Fig. 5). XH-7 was fine mapped (data not shown) and demonstrated an almost identical restriction fragment pattern to WEE strain BFS 1703 (Hahn *et al*, 1988).

DNA SEQUENCE ANALYSIS OF THE 26S REGION OF WEE

XH-7 was completely DNA sequenced using a combination of manual (see Fig. 6) and automated sequencing (see Fig. 7). The manual sequencing was very time consuming when compared with the ABI (Foster City, CA) automated sequencer. Manual sequences were analysed on the Speed Reader^{TM} (Intelligenetics, Mountain View, CA). While being an improvement over the visual reading of autoradiographs of DNA sequencing gels, it is still a

slow and tedious process as compared to automated fluorescent systems. A complete DNA sequence was obtained for clone XH-7 (see Appendix A) using the primers listed in Table 1. In the nucleotide sequence, 63 nt changes were detected , with only 11 causing amino acid substitutions from WEE BFS 1703 (see Fig. 8). The 63 nt changes were generally scattered throughout the 26S gene segment as follows: 5' noncoding region - 1; NC - 16; E3 - 2; E2 -15; 6K - 1; E1 - 19, 3' noncoding region - 9. The 26S promoter (101-124) and the tandem repeats (3896-3935, 3954-3993) in the 3' noncoding region were conserved (1 nt change to make it an exact repeat in 71V-1658 - Appendix A). The XH-7 construct had a continous open reading frame beginning and terminating at the same amino acid residues as strain BFS 1703 (Appendix B). The amino acid sequence was highly conserved, with 10 of the 11 amino acid differences being conservative changes. Only the Cys to Arg (28) in the NC protein would be considered nonconserved (see Fig. 8). The XH-7 clone was truncated at the 3' end of the genome with respect to the BFS 1703 sequence. This is likely a deletion, as restriction analysis of the pcDW-12 as compared with pcDW-6 and 7, showed a smaller band mapping to the 3' end (1600 bp instead of 1700 bp for pcDW-6 and 7). In addition, the poly A tract at the 3' end was completely missing. To confirm this, the 3' ends of both pcDW-6 and pcDW-7 could be sequenced. Further full-length cDNA constructs to WEE strain 71V-1658 were made using primers WEE5'Sst and WEE3'HindIII in a RT-PCR reaction. A set of XTRAH clones were isolated, and XTRAH-5 was characterized by restriction mapping (data not shown) as containing the entire 26S coding sequence.

ANALYSIS OF THE 5' WEE NONSTRUCTURAL REGION

One of the approaches initially examined for vaccine research was to obtain an infectious clone for mutagenesis studies. In order to accomplish this task, the entire genome of WEE must be cloned. Sequence information was available to the middle of the nonstructural region (Weaver *et al*, 1993), however, no sequence information was available for the 5' end of the nonstructural region. Nevertheless, it was deduced that the 5' termini from a number of alphaviruses were highly conserved (Ou *et al*,1983), and this information was used to devise a cloning strategy for the region (see Fig. 2). Two primers were made to the 5' termini, one based on the SIN sequence (SIN5') and the other based on a combination of the EEE and HJ sequences (EHJ5').

RT-PCR reactions were carried out using three sets of PCR primers as indicated in Fig. 2, to generate 3 overlapping fragments for reconstruction of the WEE nonstructural region. For amplification of the 5' nonstructural gene fragment (WNS-A), the EHJ5' primer worked well (see Fig. 9), generating a band of the expected MW. This is in contrast to the SIN5' primer which generated only a multitude of smaller sized bands, of which many were eliminated when Perfect Match^{TM} was added to the reaction mix (see Fig. 9). In general, Perfect Match^{TM} increased the stringency of the PCR reaction, eliminating some extraneous bands. Following gel purification and ligation of the WNS-A, WNS-B and WNS-C fragments, transformant colonies were identified by colony hybridization, and rapid plasmid preparations were restriction analyzed (data not shown). Isolates of the proper size were identified for WNS-A (WNSTA-3, 4, 5 & 6) and WNS-C (WNSTC-1, 2, 3 & 4) fragments. These were sequenced with the T7 and M13 reverse primers from the pCRII multicloning region.

The clustal method of homology comparison was used for both DNA and protein sequences in the Lasergene^{TM} program. The WNSTA-3 5' DNA sequence showed high homology to HJ (Ou *et al,* 1983), EEE North American (NA) (Weaver *et al*, 1994) and EEE South American (SA) (Ou *et al,* 1983) as illustrated in Fig. 10a and Appendix C. Homology to SIN strain HR was more distant (data not shown), as illustrated in the phylogenetic tree (see Fig. 10b). The First 19 nt of WNSTA-3 are derived from the EHJ5' primer and were not included in the homology calculations, as they may not correspond exactly to the WEE 71V-1658 5' end of the genome. If the corresponding, translated open reading frames of WNSTA-3, HJ, EEE-NA, EEE-SA and SIN are compared, fairly high homology is obtained for WEE, HJ, EEE-NA and EEE-SA, as shown in Fig. 11a. Structures which were highly conserved in all alphaviruses, such as a 51 nt sequence (nt138-188) and a 22 amino acid sequence (residues 34-55) (Ou *et al*, 1983) were also present in the WEE sequence. The 3' end of WNSTA-3 was sequenced for 233 nt, and a comparison to the BFS-1703 sequence (Weaver *et al*, 1994) using the Clustal method yielded very high homology.

DISCUSSION

When vaccine research on WEE was formulated at the start of the project, Highlands J strain B320 was included, as its crossreactivity and reduced virulence in humans made it a good vaccine candidate. Previous studies had postulated that HJ was a variant of WEE based on its antigenic cross-reactivity (Hayes and Wallis, 1977; Karabatsos *et al*, 1988) and it was referred to as the eastern strain of WEE virus in the early literature. A cDNA library to HJ strain B320 was produced containing a number of isolates whose inserts were reactive with the HJ probe (see Table 2). However, with the limited sequence information available and the priority to produce a WEE specific subunit vaccine, the research focus was shifted to WEE. Nevertheless, in order to glean further information on the pathogenesis of WEE, it would be of interest to define how HJ virus is related, through DNA sequence analysis.

The WEE library was disappointing in that a full-length clone of the 26S region could not be obtained. The largest clones were estimated to be 2900 and 3100 bp respectively, for pcDW-6 and pcDW-12. The longer clone, pcDW-12 was chosen to serve as the starting material for engineering of the full-length 26S region (Fig. 1). Utilizing the advances in PCR cloning, the 5' end of phTW5'-3 was fused at the *NcoI* site to the bulk of pcDW-12 to form the XH-7 clone. A set of monoclonal antibodies to WEE strain B11 have been made and partially characterized by ELISA and western blot analysis (unpublished results). These were reactive to inactivated WEE 71V-1658 in western blots, and will be valuable in defining levels of protein produced *in vitro* (tissue culture, baculovirus, vaccinia) or *in vivo* (DNA immunization). The restriction map of 71V-1658 was highly conserved when compared to BFS 1703 and it could be used to restriction map isolates and directional clone (cloning using two different restriction sites) PCR fragments.

One technological change which had an immense impact on the outcome of the project was the switch in DNA sequencing from a manual method to an automated fluorescent method. The amount of sequence data that can be generated by the fluorescent sequencers is astounding. A year's worth of work could be easily cut down to one or two months. The need for radioistope was eliminated and samples were easier to set up, run and analyze. The raw sequence data

(ABI chromatograms) could be directly imported into the Power Macintosh-based Lasergene^{TM} software for evaluation.

The WEE virus strain 71V-1658 was selected by Dr Trent as a vaccine isolate, as it was a relatively recent isolate of WEE. The surprisingly high degree of homology between the two strains of WEE is extended to the nucleotide sequence, as shown by homology comparison in DNASTAR to the BFS 1703 sequence (see appendix A). The majority of the 63 nt changes were silent mutations. Out of eleven amino acid changes in the polyprotein, 10 produced conservative amino acid changes, and only one at residue 28 (Cys → Arg) could be considered radical. The viruses were isolated around the same time period, however, this high degree of conservation was also noted with a South American isolate, WEE strain CBA (personal communication - M. Parker, USAMRIID). In addition, the N-terminal regions of the E1 and E2 proteins were also conserved when compared to the McMillan strain of WEE which was isolated in 1957 (Bell *et al*, 1983). From over 60 residues which were determined for the N-termini of each of E1 and E2 proteins of the McMillan strain, only 3 amino acid changes were noted for 71V-1658 (see Fig. 8). With these indications that the amino acid sequence is highly conserved amongst WEE strains, one strain of WEE should be able to provide protection to most of the WEE isolates. While the 3' noncoding region of XH-7 was deleted (21 nt), it should not affect the expression of the structural genes, especially in constructs which contain a poly-A tailing and/or a transcription terminator signal.

The 5' end of the WEE genome was surprisingly well conserved when compared with the 5' end of the HJ or EEE genomes (see Fig. 10a). It was previously hypothesized that WEE was a recombinant virus, derived from EEE and a SIN -like virus, with the nonstructural region and capsid protein derived from EEE (Hahn *et al*, 1988). This data confirms that the 5' end of WEE is derived from EEE and not from a SIN ancestor (see Fig. 10b and 11b). The limited HJ virus sequence was also highly conserved in this region of the genome, however, more sequence data for HJ will have to be obtained to draw definitive conclusions. The 2.2 kb WNS-A fragment represents a part of the WEE genome that has not been examined to date. In the course of the research, an infectious clone approach had been dropped in favor of a subunit vaccine approach

(DNA immunization, baculovirus or yeast expression). However, it would be of interest to complete the sequencing of the WNS-A fragment from a scientific point of view.

CONCLUSION

A full-length clone for the 26S region of WEE strain 71V-1658 has been obtained and completely sequenced. Only eleven of the sixty-three nucleotide changes resulted in amino acid changes when the sequence was compared to WEE strain BFS-1703. The high degree of conservation of the structural proteins was maintained when compared to the N-terminal regions of the E1 and E2 envelope proteins of theWEE McMillan strain. The conserved nature of the structural proteins would indicate that one strain of WEE should be able to cross-protect against all WEE strains. Furthermore, the 5' end of the genome was cloned as a 2.2 kb fragment and sequenced for 247 nt on 5' end and 233 nt on 3' end of the fragment (WNS-A). It demonstrated high homology to both HJ and EEE, adding further evidence that the entire 5' nonstructural region of WEE was derived from EEE. This data also demonstrates the close relationship between HJ and WEE at the genetic level, which was previously implied by antigenic cross-reactivity. In summary, the genetic cloning and sequencing of the WEE 26S region marks the critical first step in the generation of a subunit vaccine to WEE.

Acknowledgements:

The authors wish to thank Dennis Trent and Richard Kinney (CDC, Fort Collins, CO) for the kind gift of WEE and HJ RNA. Jonathan Smith, George Ludwig, Mike Parker, Mike Shoemaker, Pamela Glass and Jerry Jennings of the Virology Division, USAMRIID for the use of BL-3 facilities, help and continuing support with the project. Shelly Ewing, Randy Lynde and Vicki Roberts for photographs and illustration.

REFERENCES

Ausubel, F.M. *et al,* editors (1995). *Current Protocols in Molecular Biology*, John Wiley & Sons, New York.

Bell, J.R., M.W. Bond, M.W. Hunkapiller, E.G. Strauss, J.H. Strauss, K. Yamamoto and B. Simizu (1983). Structural proteins of western equine encephalitis virus: amino acid compositions and N-terminal sequences. *J. Virol.* **45**:708-714.

Bird, B.R. and F.T. Forrester (1981). *Basic Laboratory Techniques in Cell Culture*, U.S. Department of Health and Human Services, Centers for Disease Control, Atlanta, Georgia.

Birnboim, H.C., and J. Doly (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**:1513-1523.

Calisher, C.H. (1994). Medically important arboviruses of the United States and Canada. *Clin. Microbiol. Rev.* **7**:89-116.

Calisher, C.H., N. Karabatsos, J.S. Lazuick, T.P. Monath and K.L. Wolff (1988). Reevaluation of the western equine encephalitis antigenic complex of alphaviruses (family Togaviridae) as determined by neutralization tests. *Am. J. Trop. Med. Hyg.* **38**:447-452.

Danner, D.B. (1982). Recovery of DNA fragments from gels by transfer to DEAE paper in an electrophoresis chamber. *Anal Biochem* **125**:139-142.

Davis, R.W., D. Botstein and J.R. Roth (1980). *A Manual for Genetic Engineering: Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23**:641-651.

England, J.A., G.N. Breningstall, L.J. Heck, J.S. Lazuick, N. Karabatsos, C.H. Calisher and T.F. Tsai (1986). Diagnosis of western equine encephalitis in an infant by brain biopsy. *Pediatric Infect. Disease* **5**:382-384.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557-580.

Hahn, C.S., S. Lustig, E.G. Strauss and J.H. Strauss (1988). Western equine encephalitis virus is a recombinant virus. *Proc. Natl. Acad. Sci. USA* **85**:5997-6001.

Harlow, E. and D. Lane (1988). *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Hayes, C.G. and R.C. Wallis (1977). Ecology of western equine encephalomyelitis in the eastern United States. *Adv. Virus Res* **21**:37-83.

Holmes, D.S. and M. Quigley (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem* **114**:193-197.

Horne, D., T. Jadavji, M. Lemay and H. Rabin (1989). A case of western equine encephalitis - Alberta. *Can. Diseases Weekly Report* **15**:256-257.

Karabatsos, N., A.L. Lewis, C.H. Calisher, A. Hunt and J. T. Roehrig (1988). Identification of Highlands J virus from a Florida horse. *Am J. Trop. Med. Hyg.* **39**:603-606.

Levine, B., H.M. Hardwick, B.D. Trapp, T.O. Crawford, R.C. Bollinger and D.E. Griffin (1991). Antibody-mediated clearance of alphavirus infection from neurons. *Science* **254**:856-860.

Lupton, H.W. and G.F. Meadors (1989). Immunization of persons at-risk of exposure to western equine encephalitis virus: Continued assessment of the safety and effectiveness of western equine encephalitis vaccine, inactivated, TSI-GSD 210, Lot 1-81, IND - 2013, USAMRIID.

McConaughy, B.L., C.D. Laird and B.J. McCarthy (1969). Nucleic acid reassociation in formamide. *Biochemistry* **8**:3289-3295.

Office of Biosafety, Laboratory Centre for Disease Control, Health Protection Branch, Canada (1989). Material Safety Data Sheet - Eastern and western equine encephalitis.

Ou, J-H., E.G. Strauss and J.H. Strauss (1981). Comparitive studies of the 3'-terminal sequences of several alphavirus RNAs. *Virology* **109**:281-289.

Ou, J-H., E.G. Strauss and J.H. Strauss (1983). The 5' terminal sequences of the genomic RNA of several alphaviruses. *J. Mol Biol.* **168**:1-15.

Peters, C.J. and J.M. Dalrymple (1990). Alphaviruses. *In Fields Virology, 2nd Edition*, B.N. Fields, *et al* (eds.), Raven Press, New York, p. 713-761.

Rennels, M.B. (1984). Arthropod-borne virus infections of the central nervous system. *Neurol. Clin.* **2**:241-254.

Roehrig, J.T. (1986). The use of monoclonal antibodies in studies of the structural proteins of togaviruses and flaviviruses. *In The Togaviridae and Flaviviridae*, S. Schlesinger and M.J. Schlesinger (eds.). Plenum Publishing Corp., NY, p. 251-278.

Sambrook, J., E.F. Fritsch and T. Maniatis (1989). *Molecular Cloning: A Laboratory Manual* (2nd Ed.), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Saravolac, E.G., B. Kournikakis, L. Gorton and J.P. Wong (1996). Effect of liposome-encapsulation on immunomodulating and antiviral activities of interferon-γ. *Antivirol. Res.* **29**:199-207.

Strauss, E.G., C.M. Rice and J.H. Strauss (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133**:92-110.

Strauss, J.H. and E.G. Strauss (1994). The alphaviruses: Gene expression, replication and evolution. *Microbiol. Rev.* **58**:491-562.

Weaver, S.C., *et al* (1993). A comparison of the nucleotide sequences of eastern and western equine encephalomyelitis viruses with others in the alphavirus-like superfamily. *Virology* **197**:375-390.

Wong, J.P., J.W. Cherwonogrodzky, V.L. Di Ninno, L. Stadnyk and M.H. Knodel (1992). Liposome potentiation of humoral immune response to lipopolysaccharide and o-polysaccharide antigens of *Brucella abortus*. *Immunology* **77**:123-128.

Table 1: WEE 26S REGION PRIMERS

✓	•	Name	Length	Sequence	Note
✓		HJ-3'	30	GCATTATGCACTGTGCTTCCTCGGAAGTAC	Highlands J 3' end
✓		WEEPRO	30	AATCACCCTCTACGGCTGACCTAAATAGGT	WEE 26S 96-125, + STRAND PROMOTOR
✓		WEEPR-SST	24	GGCTGAGCTCAATAGGTGACGTAG	WEE 26S 109-132, + STRAND SstI SITE
✓	•	WEE3'	30	GTAGTGTATATTAGAGACCCATAGTGAGTC	WEE 26S 4086-4057, - STRAND
✓		WEE5'Sst	20	TCCAGATACGAGCTCATACT	WEE 26S 1-20, + STRAND SstI SITE
✓		WEEN1	20	GGTGCCGCTGGAGGCCGTTT	WEE 26S 238-257, + STRAND
✓		WEEN1A	20	GATCTTAGGAGGTCGATAGC	WEE 26S 285-304, + STRAND
✓	•	WEEN2	20	GGCTGATGAAACCACTCCAC	WEE 26S 547-566, + STRAND
✓	•	WEEN3	20	CCACCCGTGTGCTATTCACT	WEE 26S 990-1009, + STRAND
✓	•	WEEN3A	20	CGCCGTGTTTCAGCCCAATA	WEE 26S 990-1009, + STRAND
✓	•	WEEN4	20	TCACGAGCGGAGCATCTGAG	WEE 26S 1459-1478, + STRAND
✓	•	WEEN5	20	GGCATCACCCTCCACCTGAC	WEE 26S 1959-1978, + STRAND
✓	•	WEEN6	20	TTGTTATTCTGTTCCGCTGC	WEE 26S 2469-2488, + STRAND
✓	•	WEEN7	20	CTATTGATCATGCAGTCGCA	WEE 26S 2893-2912, + STRAND
✓	•	WEEN8	20	AGTGGAGCCTCTGCGAGCGT	WEE 26S 3335-3354, + STRAND
✓	•	WEEN9	20	GAGGAGTGGGCGGGAAAGGC	WEE 26S 763-782, + STRAND
✓	•	WEEN10	20	CTAAAACTCGATGTATTTCC	WEE 26S 3952-3971, + STRAND
✓	•	WEEN11	20	ACGCGAACGAAGATGAACGG	WEE 26S 1752-1771, + STRAND
✓	•	WEEN12	20	ACTGTCATTGTGCTGTGTGG	WEE 26S 2214-2233, + STRAND
✓	•	WEEN13	20	CACAGTCATTCCTTCACCAC	WEE 26S 2705-2724, + STRAND
✓	•	WEEN14	20	CGTCATCAGAAAGGGGCTTG	WEE 26S 3080-3099, + STRAND
✓	•	WEEN15	20	CAAAGCTGACAGGGAGGGAC	WEE 26S 3509-3528, + STRAND
✓	•	WEEN16	20	GGAAAGCTGGTAAAGTGCCA	WEE 26S 1554-1573, + STRAND
✓		WEEN0	20	GGAGAACCACATAAAGTCGA	WEE 26S 3714-3733, + STRAND
✓		WNSP1	25	GGCTAACGTGGACAGGGACGTGATG	WEE 26S 57-33, - STRAND
✓	•	WEEP0	20	GGCTATCGACCTCCTAAGAT	WEE 26S 305-286, - STRAND
✓	•	WEEP0A	20	CTGTCGGTTCCCTGGTTTAG	WEE 26S 458-439 - STRAND
✓	•	WEEP1	20	CTGGGGAACGTCGCCATACT	WEE 26S 656-637, - STRAND
✓	•	WEEP2	20	CGTTCTCCAGCAGCGTGTCG	WEE 26S 1083-1064, - STRAND
✓	•	WEEP2A	20	TATTGGGCTGAAACACGGCG	WEE 26S 1202-1183, - STRAND
✓		WEEP3	20	CTTCAAGTGATCGTAAACGT	WEE 26S 1592-1573, - STRAND
✓	•	WEEP4	20	ACTCCAGCCCTTCTCGCCCC	WEE 26S 2097-2078, - STRAND
✓		WEEP5	20	GTTCGACCAACGCCTTATAC	WEE 26S 2610-2591, - STRAND
✓	•	WEEP6	20	AAGGGTGAAAAAGCGGCTGA	WEE 26S 3067-3048, - STRAND
✓		WEEP7	20	GGTGATTCTGATGATCTCAC	WEE 26S 3430-3411, - STRAND
✓	•	WEEP8	20	TGGAAACTGCCGCCTGGAAT	WEE 26S 3759-3740, - STRAND
✓	•	WEEP10	20	CCTTGATGTCATGGTCGTGG	WEE 26S 1344-1325, - STRAND
✓	•	WEEP11	20	TGCACTGAGTGGTCTGTGTG	WEE 26S 1870-1851, - STRAND
✓	•	WEEP12	20	ATGTTTCAGCGTTGGTTGGC	WEE 26S 2391-2372, - STRAND
✓		WEEP13	20	GTGTTCTCACTGTCACAGAA	WEE 26S 2851-2832, - STRAND
✓	•	WEEP14	20	ATGTGTGGTCGCTTCCTTCA	WEE 26S 3584-3565, - STRAND

Table 2: SUMARY OF HJ ISOLATES

Isolate	HJ-3'	WEE	Size of Inserts
A13	- -+	- - -	1400, 500, 450
A14	w - -	- - -	ND
A15	+ - -	- - -	ND
A16	w - -	w - -	500
A17	- w -	w - -	ND
A18	ww -	w - -	ND
A19	w + +	w - -	1000
A20	w - -	w - -	ND
A21	w - -	w - -	ND
A22	ww +	w - -	1200
A23	- w +	- - -	500
A24	- + +	- - -	500, 450
A25	- - -	- - -	300
A26	w - -	- - -	3K?, 500, 300
A27	ww +	- - -	600
A28	w - -	- - -	3k?, 600
A29	w -w	- - -	1400, 600
A30	- w +	- - -	1000
A31	- - -	w - -	500
A32	- - -	w - -	1400
A33	w -w	w - -	1400
A34	- w +	w - -	400
A35	ww +	+ - -	1600, 400
A36	- w +	- - -	400
B13	- - +	- - -	1000, 400
B14	- - -	- - -	450
B15	w - -	- - -	ND
B16	+ - -	- - -	ND
B17	w -w	- - -	600?, 500
B18	ww +	- - -	600
B19	+ - -	w - -	ND
B20	- w +	w - -	600
B21	w - +	w - -	500
B22	w -w	w - -	500
B23	+ w +	+ - -	550, 400

Isolate	HJ-3'	WEE	Size of Inserts
B24	w + +	+ - -	500
C1	- - -	- - -	ND
C2	- - -	- - -	1700
C3	- - -	- - -	300
C4	- - -	- - -	450
C5	w - -	w - -	650
C6	w -w	w - -	450
C7	- - w	+ - -	900, 300
C8	- w w	w - -	350, 400
C9	ww +	w - -	600
C10	ww -	+ - -	600
C11	www	+ - -	1800
C12	w + +	+ - -	850
C13	- ww	- - -	1000, 300
C14	- ww	- - -	300
C15	- ww	- - -	700
C16	www	- - -	1000?, 800
C17	www	w - -	700, 600, 500
C18	- - -	w - -	600
C19	+ - -	+ - -	3K?, 600
C20	- - +	w - -	1700, 1300, 650?
C21	- - -	- - -	650, 500
D1	w + +	+ - -	1000
D2	+ w +	+ - -	600
D3	w - -	- - -	1800
D4	- - -	- - -	400
D5	w - -	- - -	400
D6	- - w	- - -	600
D7	+ - -	w - -	550
D8	+ - -	w - -	300
D9	- - w	w - -	400
D10	+ - -	w - -	450
D11	+ - -	+ - -	1100
D12	www	+ - -	1400
D13	- - -	+ - -	ND

Isolate	HJ-3'	WEE	Size of Inserts
D14	- - -	+ - -	ND
D17	- - -	- - -	400
D18	- - -	- - -	400
D19	- - -	- - -	2K, 600
D20	- ww	- - -	1000, 400
D22	- - +	- - -	900, 400?, 300
D23	- - -	- - -	950
D24	- - -	w - -	400
D25	w - -	w - -	1200, 600?
D26	- - -	- - -	1700
D27	- - -	- - -	400
D28	- - -	- - -	650
D29	+ - -	- - -	3K?, 1500
D30	w - -	- - -	1400?, 1200
D31	- w -	- - -	1200?, 800
D32	w - -	- - -	750
D33	w - +	- - -	600, 500
D34	- - w	w - -	400
D35	- - w	w - -	400
D36	- w +	w - -	500
D37	- - -	w - -	650, ?
D38	- - -	+ - -	1300, 400?
D39	- - w	- - -	1350, 400
D40	- - -	- - -	800
D41	+ ww	- - -	800, 300
D42	ww +	- - -	800, 350
D43	w -w	- - -	400
D44	- - w	- - -	400, 300
D45	- - w	w - -	1000
D46	- - w	w - -	1400
D47	- - -	w - -	350
D48	- - w	w - -	400
D49	- ww	w - -	1000
D50	- ww	w - -	1300
D51	- ww	w - -	700, 500?

HJ-3', WEE (WEEPRO) PROBES - N-LINK, 32P-HYBOND-N 32P HYBOND-C, RESPECTIVELY

SIZE - SIZE OF INSERT AFTER EXCISION WITH XHOI AND HINDIII
ND - NOT DETERMINED

VG PROSPECTS - A19, A24, A35, B23, C12, D1, D2
GOOD PROSPECTS - A13, A15, A22, A30, B13, B24, C19, C20, D11, D22, D29, D33, D41, D42

TABLE 3: SUMMARY OF WEE ISOLATES

pcDW	A	B	C	D
1	+	+	+	-
2	+	+	w	-
3	+	+	w	-
4	+	+	+	-
5	w	w	-	-
6	+	+	+	-
7	+	++	++	-
8	+	w	-	-
9	+	w	-	-
10	+	w	-	-
11	++	++	-	-
12	++	++	+	-
13	+	+	-	-
14	++	+	+	-
15	+	+	+	-
16	+	+	+	-
17	+	+	+	-
18	w	-	-	-
19	+	w	-	-
20	w	w	-	-
21	++	+	-	-
22	+	w	-	-
23	++	++	-	-
24	+	w	-	-
25	+	+	-	-
26	w	+	-	-
29	w	w	-	-
30	+	+	+	-
31	w	w	-	-
33	+	w	-	-
34	++	++	w	-
35	+	w	+	-
36	w	w	-	-
38	w	-	-	-
39	+	+	+	-

pcDW	A	B	C	D
40	+	w	w	-
41	+	+	-	-
42	+	+	-	-
43	w	w	-	-
44	+	+	-	-
45	++	+	-	-
46	+	w	-	-
47	++	+	-	-
48	+	+	-	-
49	+	+	+	-
50	w	w	w	-
53	+	+	+	-
54	+	w	-	-
55	+	w	-	-
56	w	w	-	-
57	+	w	-	-
59	+	+	w	-
61	+	w	-	-
62	+	w	-	-
63	+	w	-	-
64	w	w	-	-
67	+	w	-	-
68	w	-	-	-
69	+	w	-	-
70	w	w	-	-
71	+	w	-	-
72	w	w	-	-
73	+	+	-	-
74	+	+	++	-
75	+	w	-	-
76	w	w	-	-
77	w	w	-	-
78	+	+	-	-
79	+	W	-	-
80	-	-	-	-

pcDW	A	B	C	D
81	w	w	-	-
84	w	-	-	+
85	-	w	-	-
86	-	w	-	-
87	w	+	-	-
89	-	w	-	w
91	w	w	-	w
93	w	w	-	-
96	w	w	-	-
97	+	w	-	-
98	w	w	-	-
99	w	w	-	-
100	+	+	-	-
103	-	w	-	w
109	-	w	-	-
110	w	w	-	-
111	w	w	w	-
112	w	w	-	-
114	-	w	-	-
115	-	-	-	-
116	-	-	-	-
118	-	-	-	-
119	-	-	-	-
120	-	w	-	-
121	-	w	-	-
122	-	-	-	-
123	-	w	-	-
124	+	w	-	w
125	-	w	-	-
126	-	w	-	-
127	-	w	-	w
128	-	w	-	-

PCR Probes: A - WEEN0/WEE3' B - WEEN7/WEEP/7 C - WEEN5/WEEP5 D - WEEPRO/WEEP0
Only positive isolates included in Table.

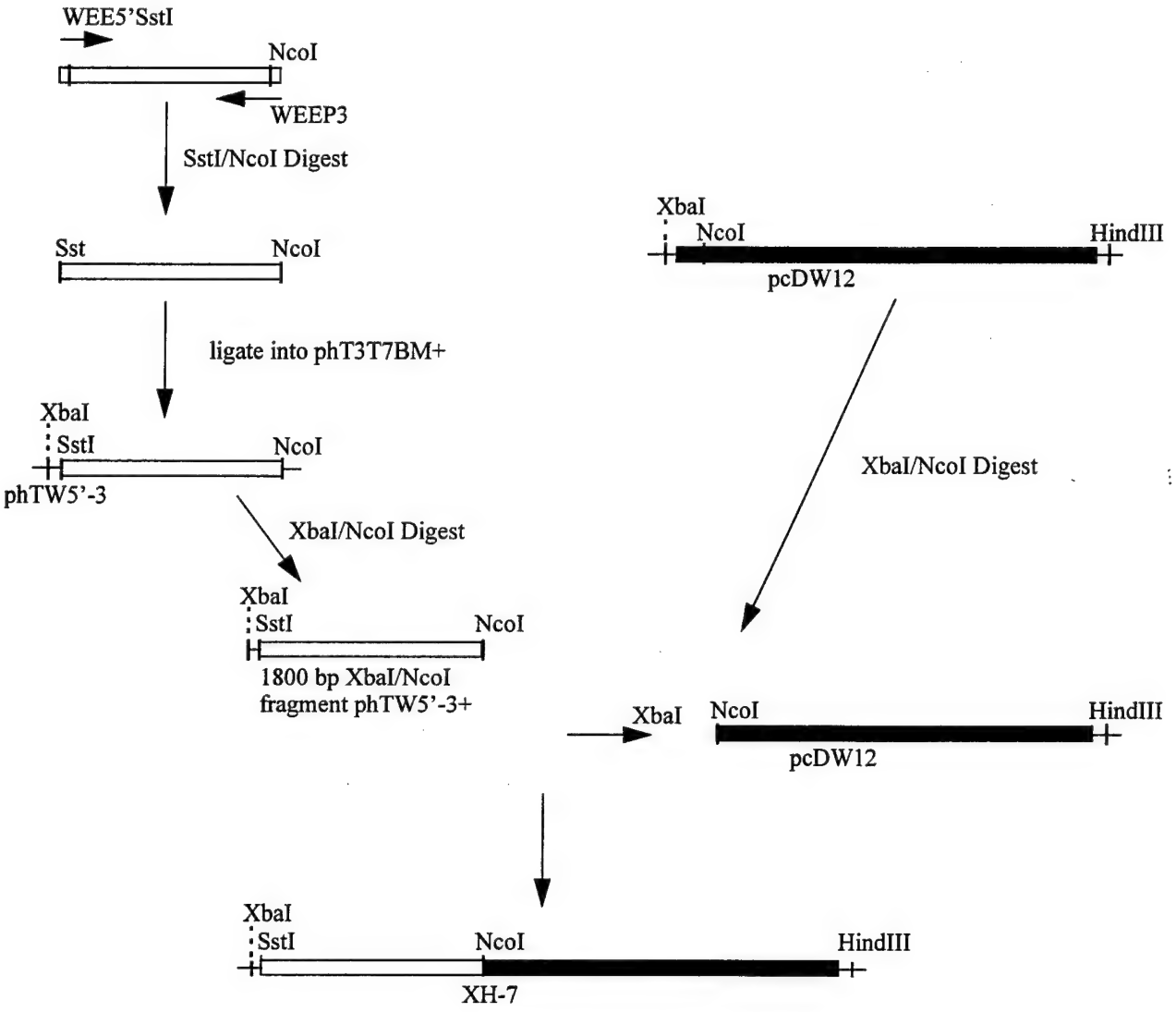
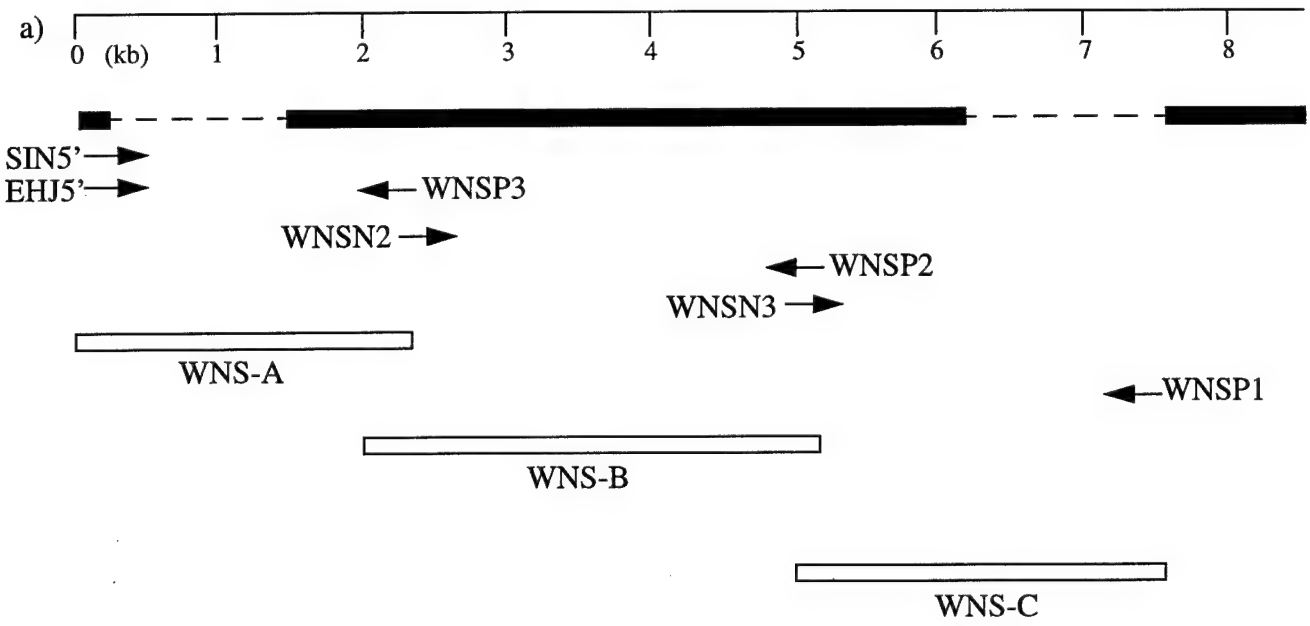


Figure 1. Construction of the Full-length WEE 26S Region

The 5' end of the 26 S region of was constructed using RT-PCR with the WEE5'Sst/WEEP3 primers as described. An 1800 bp fragment was cloned into *SstI/NcoI* sites of phT3T7BM+. The XbaI/NcoI fragment (from phTW5'-3) was cloned into the *XbaI/NcoI* sites of pcDW-12 to produce the full-length WEE 26S region isolate, XH-7.



b)	SIN5'	(1-19)	ATTGGCGGCG TAGTACACA
	EHJ5'	(1-19)	ATAGGG[C/T][A/G]TGGT[A/G]TAGAG[G/T]
	WNSN2	(811-830)	AGCAGCACCT CACAAAGTCC
	WNSN3	(3626-3645)	CGACTATGAC AGGCGGTTTG
	WNSP3	(57-33)	GGCTAACGTG GACAGGGACG TGATG
	WNSP2	(3865-3846)	GGAGCAGTAT TGAAGTCTTG
	WNSP1	(1009-990)	TAGAAGCACT GAATCAACAG

Figure 2. PCR Amplification of WEE Nonstructural Region

a) The RT-PCR cloning strategy is shown for WNS-A, B and C, together which encompass the entire nonstructural region of WEE. Primers used are indicated. b) The actual sequence of each of the primers is indicated, with the location in brackets: SIN5' and EHJ5' from Ou *et al*, 1983; WNSN2, WNSN3, WNSP1 and WNSP2 based on Weaver *et al*, 1994; and WNSP3 from Hahn, *et al*, 1988.

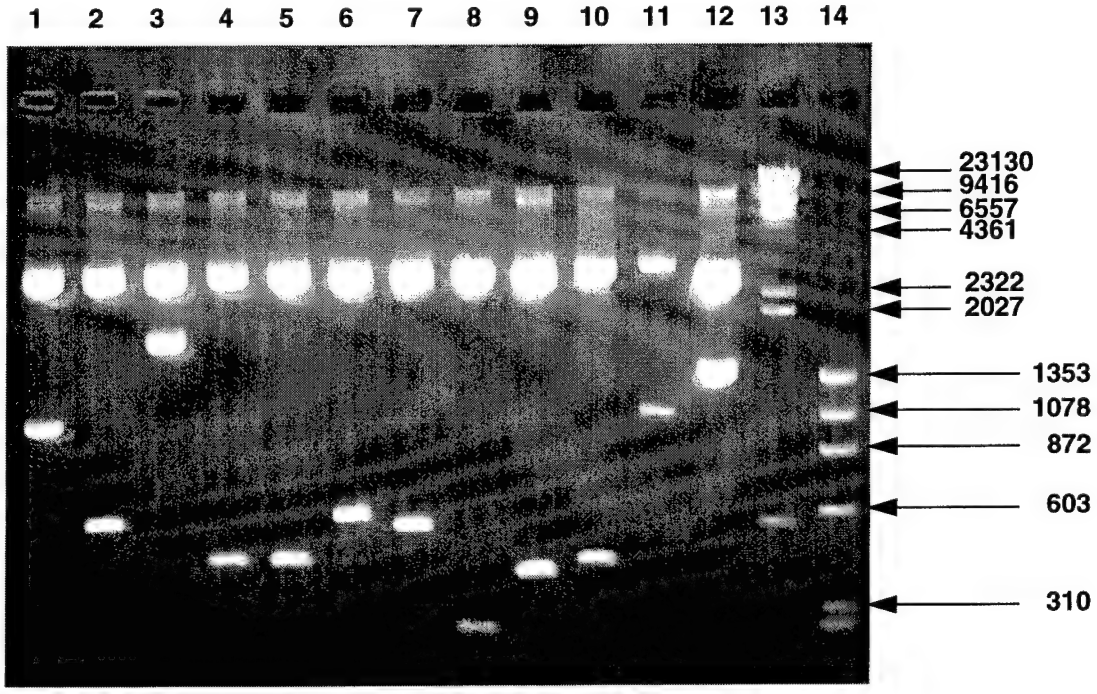


Figure 3. Restriction Analysis of HJ Isolates

Rapid plasmid preparations of HJ isolates were made up as described with the exception that a separate RNase A reaction was carried out before the final phenol/chloroform extraction and ethanol precipitation was completed. The washed and dried pellet of DNA was dissolved in 120 lof dH_2O. The restriction digest mixture consisted of 18 lDNA, 2 l10x reaction buffer (500 mM Tris-HCl, pH8.0, 100 mM $MgCl_2$, 500 mM NaCl) and 2 lof *HindIII/XhoI* (2U of each) stock diluted in 1x reaction buffer. The tubes were incubated 3 hr at 37°C. The samples were run on a 1 %(w/v) TBE agarose gel. Lanes: 1-12) HJD-1 to 12; 13) λ DNA/HindIII digested MW marker; 14) ΦX174 RF DNA/HaeIII digested MW marker.

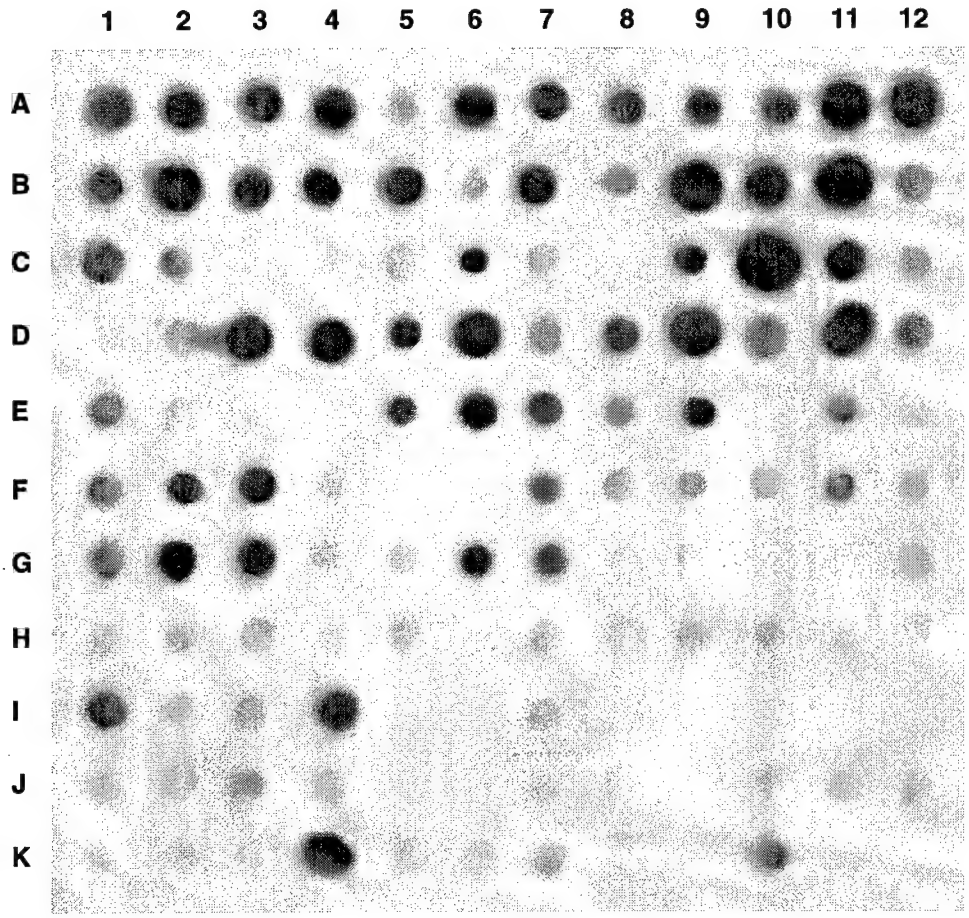


Figure 4. Dot Blot Hybridization of WEE Isolates

Rapid plasmid preparations of isolates from the WEE cDNA library (Invitrogen, San Diego, CA) were made up as described. Two 1of each sample were spotted to a nylon membrane (Hybond-N^{TM} - Amersham, Oakville, Ont.). Samples were air dried, UV fixed and hybridized with a Random primer [^{32}P]-labelled probe of the WEEN0/WEE3' PCR fragment as described. Samples: A1-12) pcDW 1-12; B1-12) pcDW 13-24; C1-12) pcDW 25-36; D1-12) pcDW37-48; E1-12) pcDW 49-60; F1-12) pcDW 61-72; G1-12) pcDW 73-84; H1-12) pcDW 85-96; I1-12) pcDW 97-108; J1-12) pcDW 109-120; K1-8) pcDW 121-128; K9) negative control (λ DNA/HindIII digested); K10) positive control (WEE 71V-1658 cDNA).

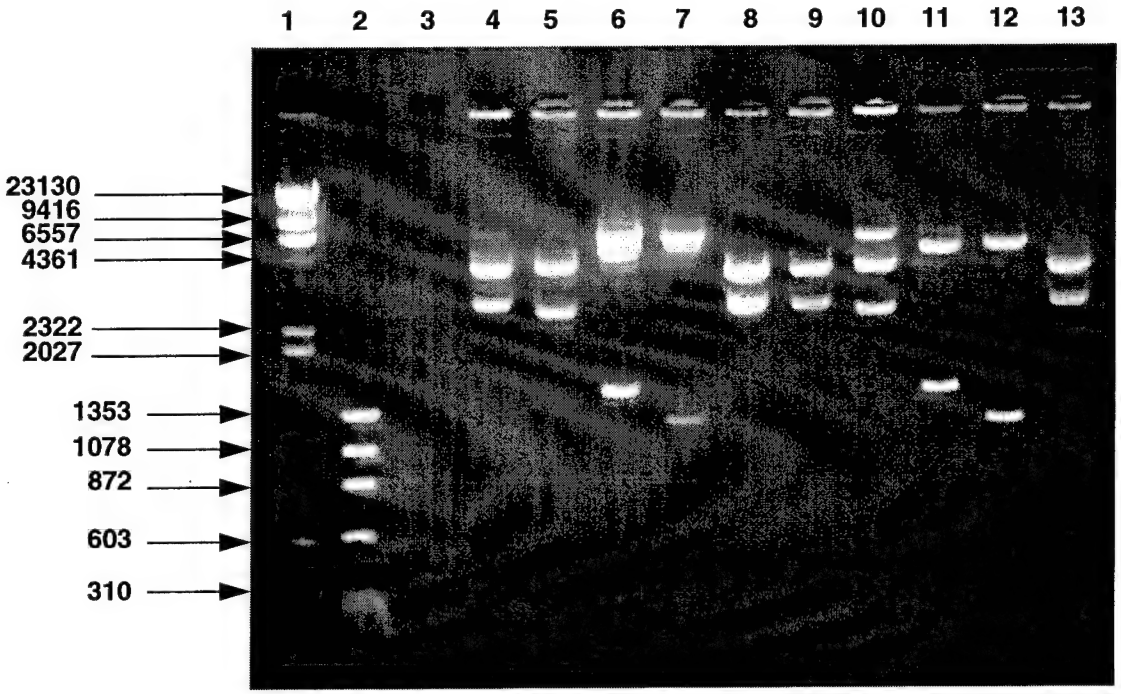


Figure 5. Restriction Mapping of WEE Isolates

Large-scale plasmid preparations were made to XH-7 and XH-12. Samples were digested with restriction enzymes as described and run on a 1 %(w/v) TBE agarose gel. Lanes: 1) λ DNA/ HindIII digested MW marker; 2) ΦX174 RF DNA/HaeIII digested MW marker. XH-7 digested with: 4) *XbaI/HindIII*; 5) *XbaI/XhoI*; 6) *XbaI/NcoI*; 7) *BglII/NcoI*; 8) *SstI*. XH-12 digested with: 9) *XbaI/HindIII*; 10) *XbaI/XhoI*; 11) *XbaI/NcoI*; 12) *BglII/NcoI*; 13) *SstI*.

a)

5'- AGGAAGTTTG TCGGTAGAGA
GGAGTACTTG TTCCCACCCG
TCCATGGAAA GCTGGTAAAG
TGCCACGTTT ACGATCACTT
GAAGGAGACG TCTGCCGGGT
ACATAACCAT GCACAGGCA
G . . . 3'

b)

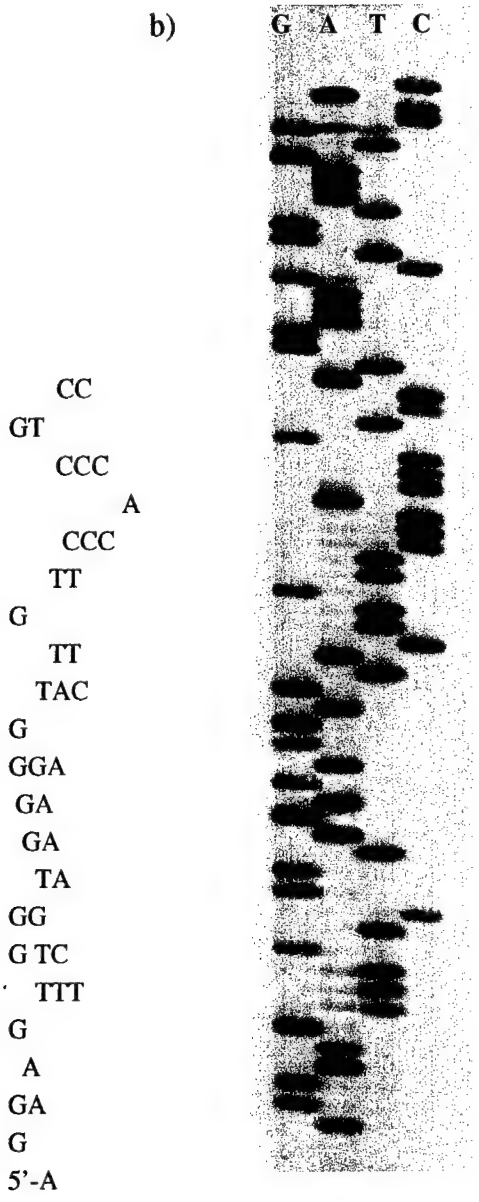


Figure 6. Manual DNA Sequence Analysis

Manual DNA sequencing was run as described, on a 8%(w/v) TBE gel (40 cm). The sequence is read sequentially from the bottom of the gel, using a set of four lanes. The sequence shown in a) used the WEEN4 primer to sequence WEE71V-1658 construct, XH-7; b) autoradiograph of the sequencing gel with the deduced sequence (5'-AGGAAGTTTGTCGGTAGAGAGAGGAG. . .)

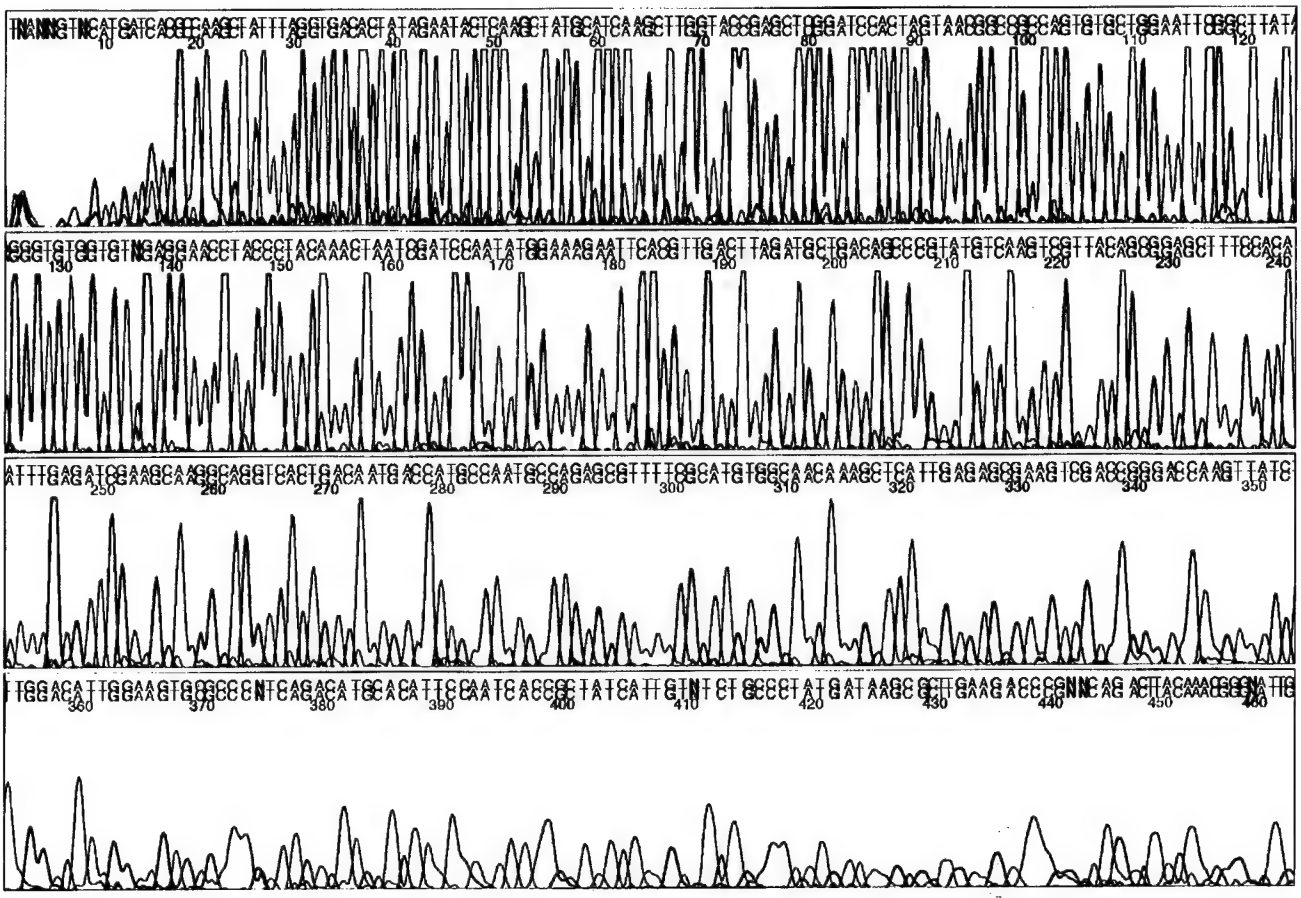


Figure 7. Fluorescent DNA Sequence Analysis

Automated sequencing was performed using a Applied Biosystem Instruments (ABI) 373 automated fluorescent sequencer, and the ABI Prism^{TM} Dye Terminator Cycle Sequencing kit (Perkin Elmer, Foster City, CA) as described in Material and Methods. A chromatogram of DNA sequence data generated from WNS-TA3 using the M13 reverse primer is shown on the following page.

Position[1]	NT Change[2]	Amino Acid Change[3]	Protein
240	T g C	Cys g Arg (28)	NC
304	C g T	Ala g Val (49)	NC
424	A g G	Lys g Arg (89)	NC
907	A g G	Lys gArg (250)	NC
1011	G g A	Ala g Thr (285)	E3
1182	G g A	Ala g Thr (342) (Ala in McMillan strain)[4]	E2
1358	T g G	Asp g Glu (400)	E2
1984	C g T	Thr g Met (609)	E2
2055	A g T	Thr g Ser (633)	E2
2698	G g A	Arg g Lys (847) (Lys in McMillan strain)[4]	E1
3669	A g T	Thr g Ser (1171)	E1

Residue 361 (a.a.) And 849 are Thr in the McMillan strain, but a Ser (361) and His (849), respectively, in 71V-1658 & BFS 1703[4].

[1] Based on the 26S sequence of WEE strain BFS 1703 (Hahn *et al*, 1988).
[2] Nucleotide change (strain BFS 1703 g 71V-1658) at the position indicated.
[3] Amino acid change (BFS 1703 g 71V-1658) at a.a. residue of 26S polyprotein ( ).
[4] Comparison with protein sequence generated to the N-terminals of the NC, E2 & E1 proteins of the McMillan strain (Bell *et al*, 1983)

Figure 8. Structural Protein Homology Between WEE Strains

The deduced amino acid sequence of the structural proteins of WEE strain 71V-1658 was compared to the deduced amino acid sequence of BFS 1703 using the Clustal method of comparison (Lasergene sequence homology programs - DNA Star, Madison, WI). Silent mutations were not included.

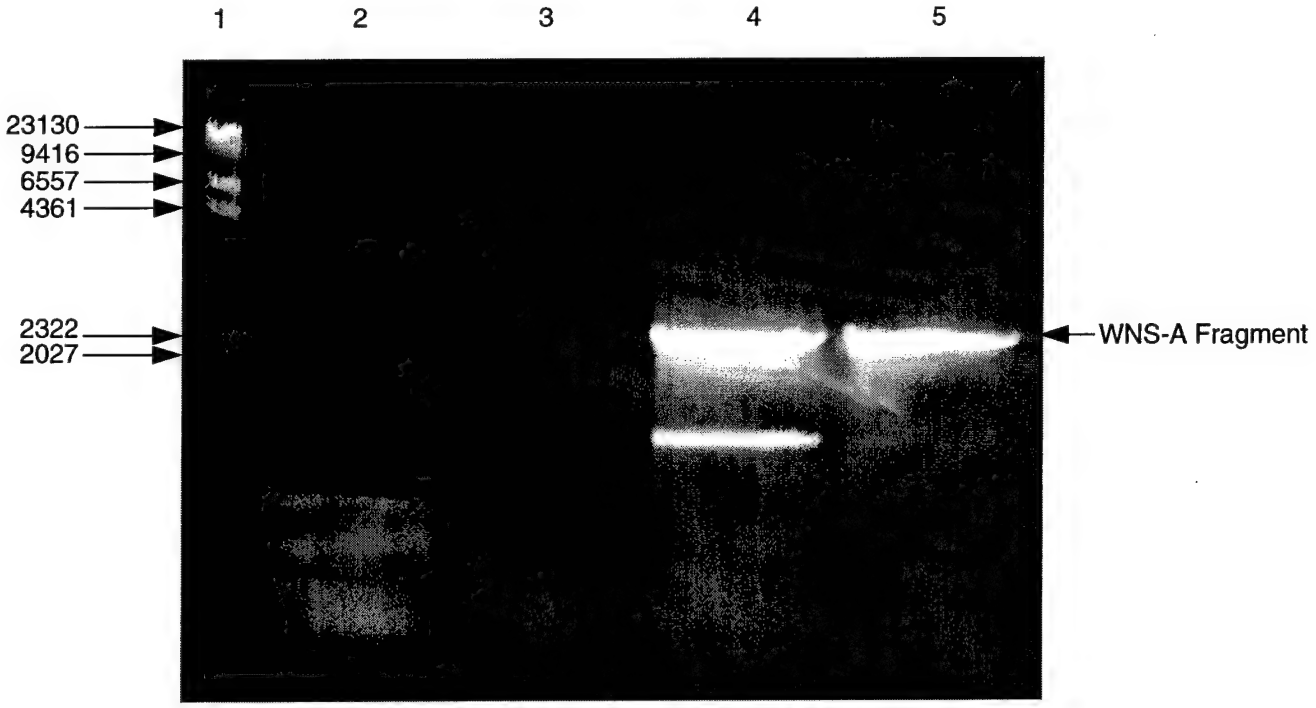


Figure 9. WEE 5' Nonstructural Region

RT-PCR reaction was run using WEE 71V-1658 RNA as template, and the following sets of primers with and without Perfect Match™ (Strategene, La Jolla, CA). a) RT-PCR reactions were run on a 1% preparatory low melting agarose gel. Lane: 1) λ *HindIII* digest; 2) SIN5'/ WNSP3; 3) SIN5'/WNSP3 with Perfect Match™; 4) EHJ5'/WNSP3; 5) EHJ5'/WNSP3 with Perfect Match™.

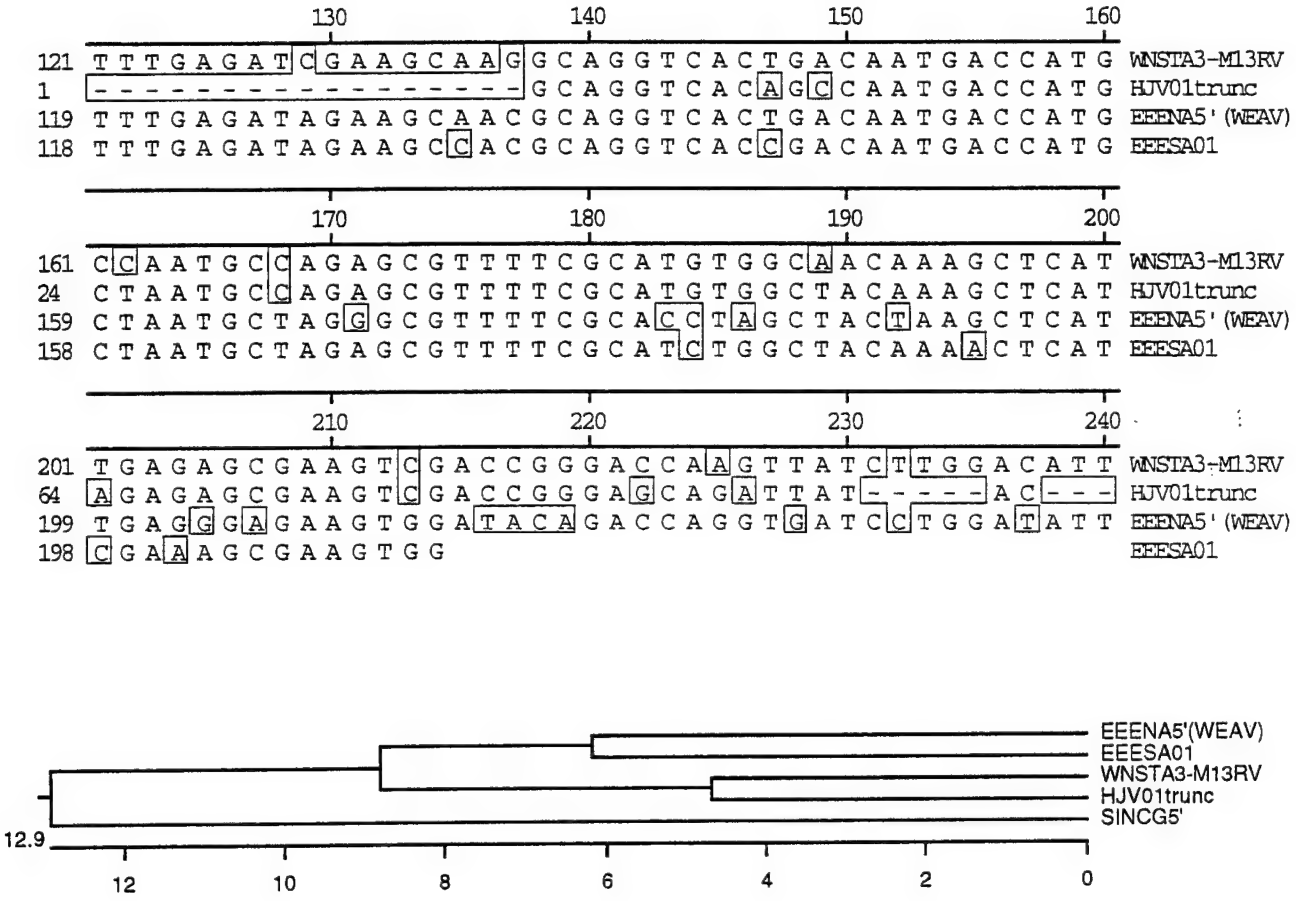


Figure 10. Nucleotide Homology In the WEE 5' Nonstructural Region

a) Sequence comparison of 5' end WEE 71V-1658 (WSNTA-3) with the HJ, EEE SA (Ou *et al*, 1983) and the EEE NA (Weaver *et al*, 1994) DNA sequences from nt 120-240, using the Clustal method; b) Phylogenetic tree of the DNA sequence comparison of 5' end of WEE, HJ, EEE SA, EEE NA and SIN HR (Strauss *et al*, 1984).

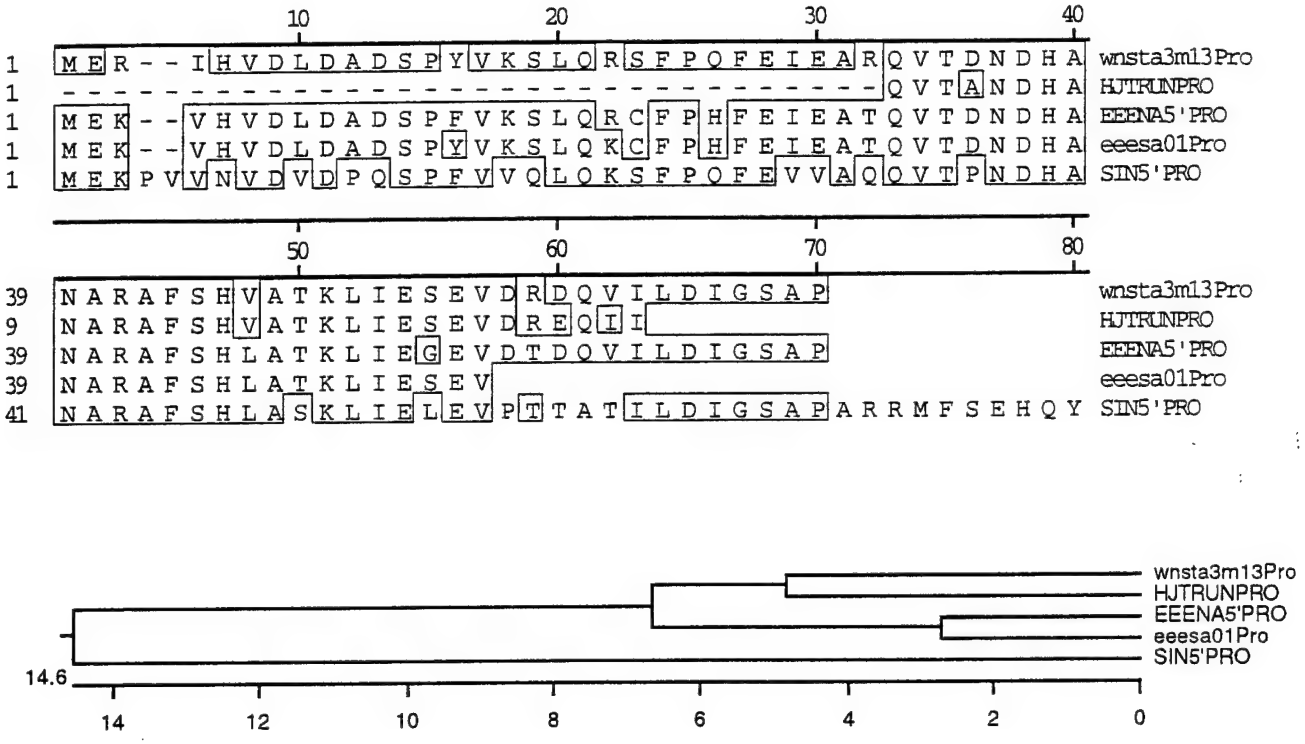
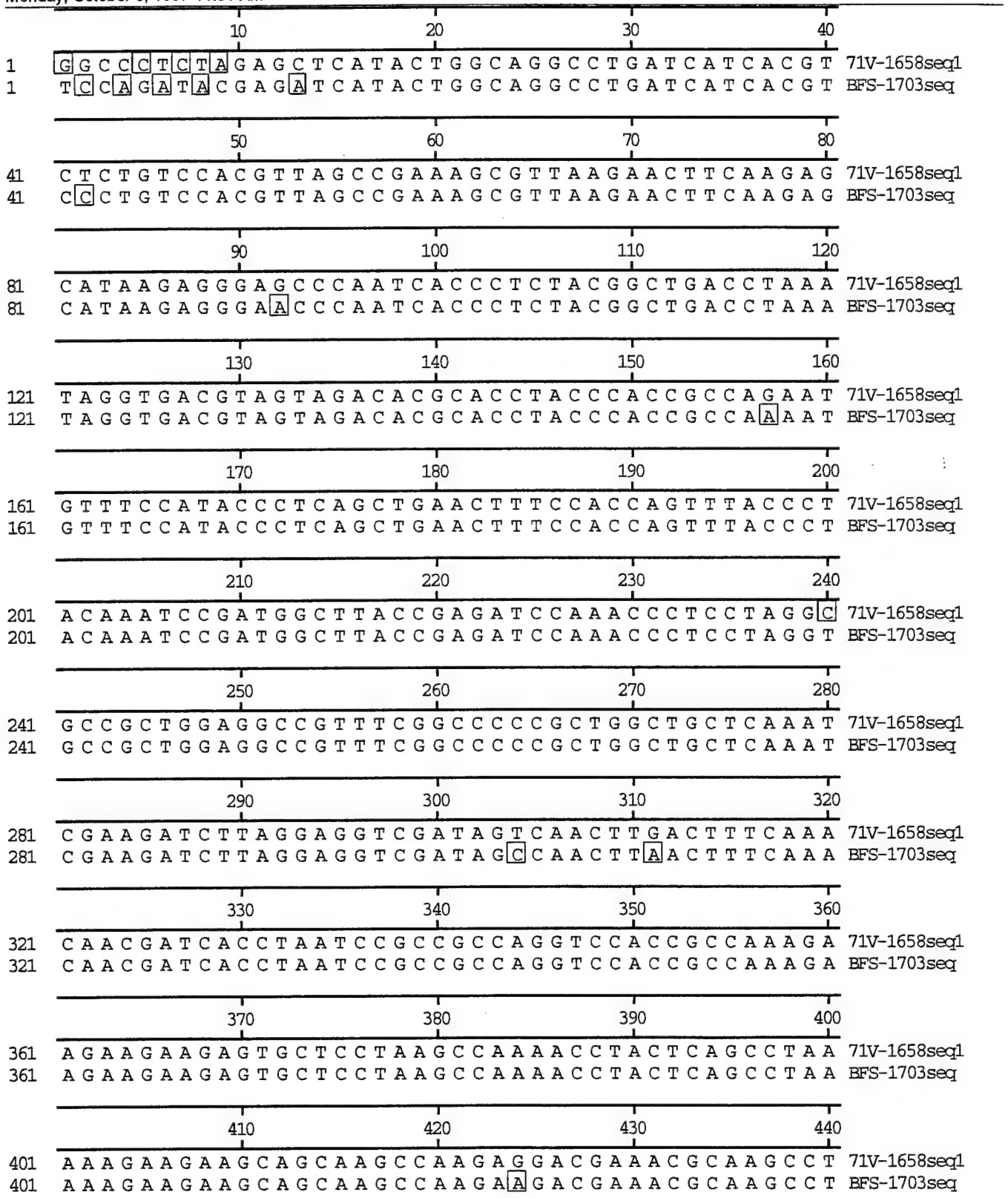


Figure 11. Homology of the N-terminus of Nonstructural Protein 1

a) Comparison of the translated protein sequence of WNSTA-3 sequence (N-terminal of nsp1) with HJ, EEE SA (Ou *et al*, 1983) and the EEE NA (Weaver *et al*, 1994) sequences using the Clustal method; b) Phylogenetic tree of the sequence comparison of the N-terminus of nsp1 of WEE, HJ, EEE SA, EEE NA and SIN HR (Strauss *et al*, 1984).

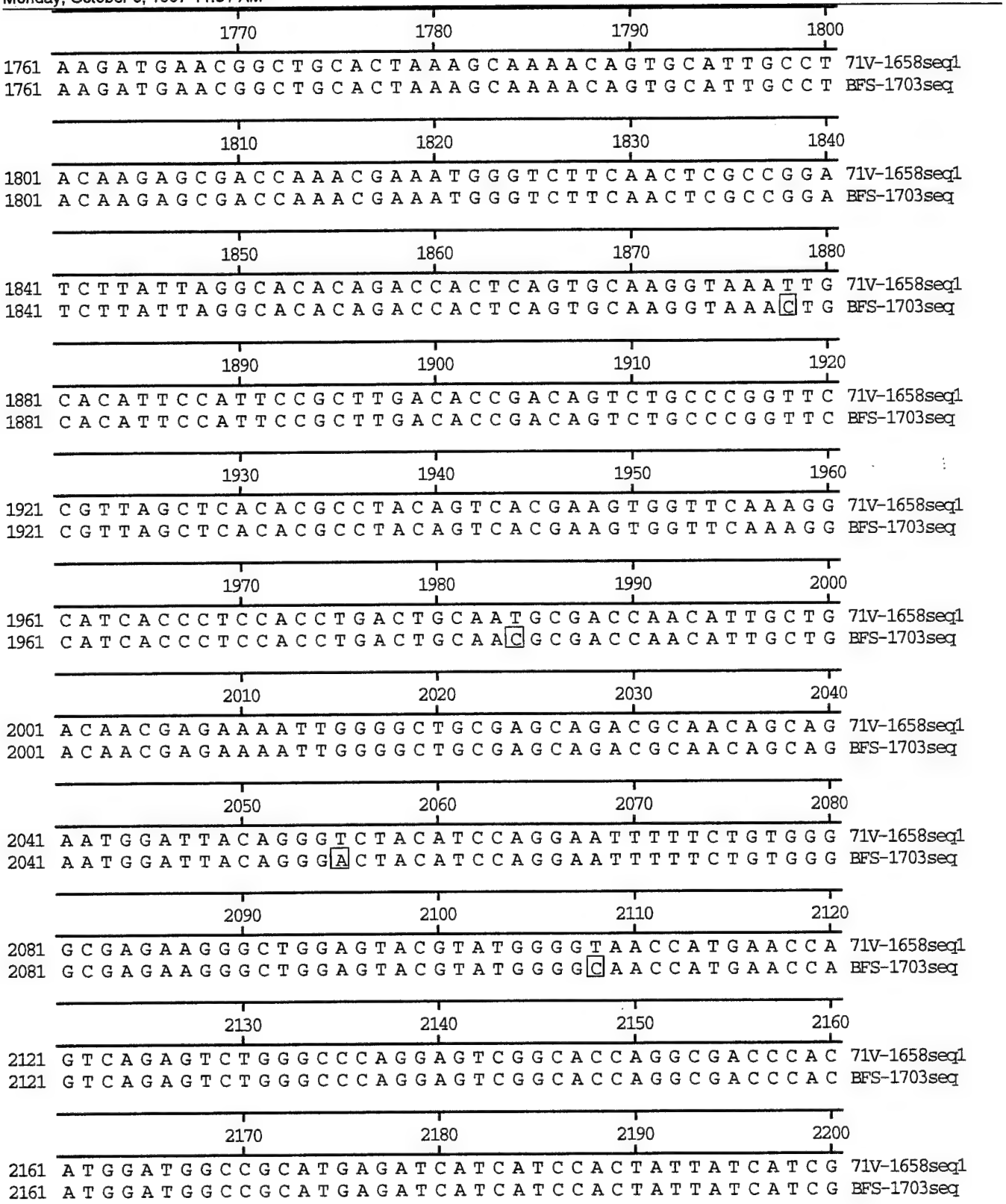
Appendix A Alignment of the 26S region of WEE strain 71V-1658 with BFS-1703 (Hahn, *et al*, 1988). The WEE sequence begins at nt 14 and ends at nt 4150. Some of the features are: the WEE 26S promoter (nt 101-124); the start of the 26S RNA (nt 120); the polyprotein start codon (nt 159); the termination codon (nt 3867); and the tandem repeats (nts 3896-3935, 3954-3993).


10 20 30 40
1 GGCCCTCTAGAGCTCATACTGGCAGGCCTGATCATCACGT 71V-1658seq1
1 TCCAGATACGAGATCATACTGGCAGGCCTGATCATCACGT BFS-1703seq
50 60 70 80
41 CTCTGTCCACGTTAGCCGAAAGCGTTAAGAACTTCAAGAG 71V-1658seq1
41 CCCTGTCCACGTTAGCCGAAAGCGTTAAGAACTTCAAGAG BFS-1703seq
90 100 110 120
81 CATAAGAGGGAGCCCAATCACCCTCTACGGCTGACCTAAA 71V-1658seq1
81 CATAAGAGGGAACCCAATCACCCTCTACGGCTGACCTAAA BFS-1703seq
130 140 150 160
121 TAGGTGACGTAGTAGACACGCACCTACCCACCGCCAGAAT 71V-1658seq1
121 TAGGTGACGTAGTAGACACGCACCTACCCACCGCCAAAAT BFS-1703seq
170 180 190 200
161 GTTTCCATACCCTCAGCTGAACTTTCCACCAGTTTACCCT 71V-1658seq1
161 GTTTCCATACCCTCAGCTGAACTTTCCACCAGTTTACCCT BFS-1703seq
210 220 230 240
201 ACAAATCCGATGGCTTACCGAGATCCAAACCCTCCTAGGC 71V-1658seq1
201 ACAAATCCGATGGCTTACCGAGATCCAAACCCTCCTAGGT BFS-1703seq
250 260 270 280
241 GCCGCTGGAGGCCGTTTCGGCCCCCGCTGGCTGCTCAAAT 71V-1658seq1
241 GCCGCTGGAGGCCGTTTCGGCCCCCGCTGGCTGCTCAAAT BFS-1703seq
290 300 310 320
281 CGAAGATCTTAGGAGGTCGATAGTCAACTTGACTTTCAAA 71V-1658seq1
281 CGAAGATCTTAGGAGGTCGATAGCCAACTTAACTTTCAAA BFS-1703seq
330 340 350 360
321 CAACGATCACCTAATCCGCCGCCAGGTCCACCGCCAAAGA 71V-1658seq1
321 CAACGATCACCTAATCCGCCGCCAGGTCCACCGCCAAAGA BFS-1703seq
370 380 390 400
361 AGAAGAAGAGTGCTCCTAAGCCAAAACCTACTCAGCCTAA 71V-1658seq1
361 AGAAGAAGAGTGCTCCTAAGCCAAAACCTACTCAGCCTAA BFS-1703seq
410 420 430 440
401 AAAGAAGAAGCAGCAAGCCAAGAGGACGAAACGCAAGCCT 71V-1658seq1
401 AAAGAAGAAGCAGCAAGCCAAGAAGACGAAACGCAAGCCT BFS-1703seq

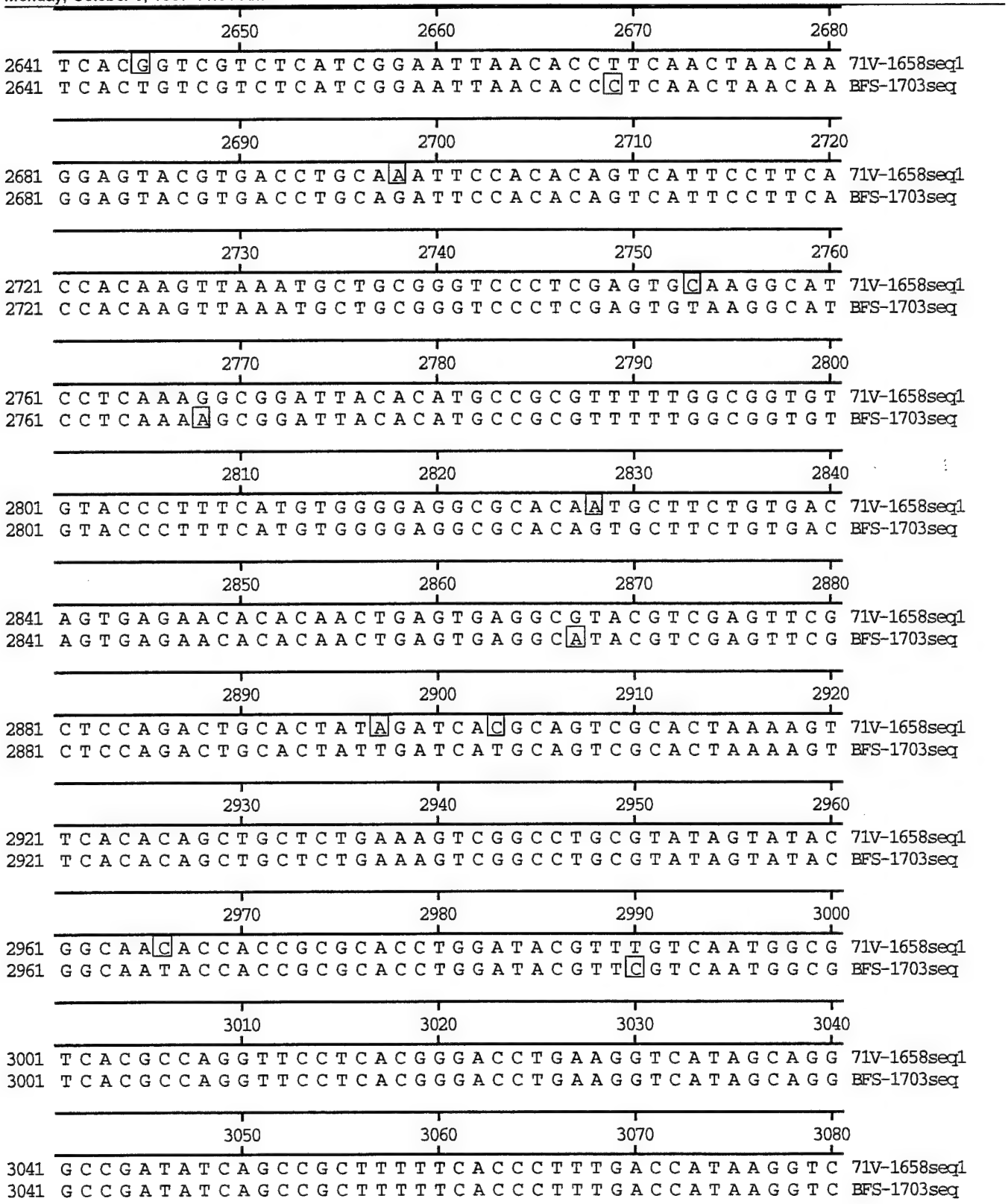
```
                    450       460       470       480
441 AAACCAGGGAAACGACAACGTATGTGTATGAAGTTGGAGT 71V-1658seq1
441 AAACCAGGGAAACGACAGCGTATGTGTATGAAGTTGGAGT BFS-1703seq

                    490       500       510       520
481 CGGACAAGACATTTCCGATCATGCTGAACGGCCAAGTGAA 71V-1658seq1
481 CGGACAAGACATTTCCGATCATGCTGAACGGCCAAGTGAA BFS-1703seq

                    530       540       550       560
521 TGGATATGCCTGCGTTGTCGGAGGAAGGCTGATGAAACCA 71V-1658seq1
521 TGGATACGCTTGCGTTGTCGGAGGAAGGCTGATGAAACCA BFS-1703seq

                    570       580       590       600
561 CTCCACGTTGAAGGAAAAATTGATAATGAGCAATTAGCGG 71V-1658seq1
561 CTCCACGTTGAAGGAAAAATCGATAATGAGCAATTAGCGG BFS-1703seq

                    610       620       630       640
601 CCGTGAAATTGAAGAAGGCTAGCATGTACGACTTGGAGTA 71V-1658seq1
601 CCGTGAAATTGAAGAAGGCTAGCATGTACGACTTGGAGTA BFS-1703seq

                    650       660       670       680
641 CGGCGACGTTCCCCAGAACATGAAATCAGACACGCTGCAG 71V-1658seq1
641 TGGCGACGTTCCCCAGAATATGAAATCAGACACGCTGCAG BFS-1703seq

                    690       700       710       720
681 TACACCAGCGACAAACCACCGGGCTTCTACAACTGGCACC 71V-1658seq1
681 TACACCAGCGACAAACCACCGGGCTTTTACAACTGGCACC BFS-1703seq

                    730       740       750       760
721 ACGGCGCAGTCCAGTATGAGAATGGGAGATTTACCGTACC 71V-1658seq1
721 ACGGCGCAGTCCAGTATGAGAATGGGAGATTCACCGTACC BFS-1703seq

                    770       780       790       800
761 GAGAGGAGTGGGCGGGAAAGGCGACAGCGGAAGACCGATC 71V-1658seq1
761 GAGAGGAGTGGGCGGGAAAGGCGACAGTGGAAGACCGATC BFS-1703seq

                    810       820       830       840
801 CTGGACAACAGAGGCAGAGTTGTGGCTATTGTTCTAGGAG 71V-1658seq1
801 CTGGACAACAGAGGCAGAGTTGTGGCTATTGTTCTAGGAG BFS-1703seq

                    850       860       870       880
841 GTGCAAATGAGGGCACGCGTACGGCGCTTTCAGTGGTCAC 71V-1658seq1
841 GTGCAAACGAGGGCACGCGTACGGCGCTTTCAGTGGTCAC BFS-1703seq
```

```
                    890        900        910        920
881 TTGGAACCAGAAAGGGGTGACCATTAGGGATACCCCCGAA 71V-1658seq1
881 TTGGAACCAGAAAGGGGTGACCATCAAGGATACCCCCGAA BFS-1703seq

                    930        940        950        960
921 GGTTCTGAACCGTGGTCACTAGTTACAGCGCTATGCGTGC 71V-1658seq1
921 GGTTCTGAACCGTGGTCACTAGTTACAGCGCTGTGCGTGC BFS-1703seq

                    970        980        990       1000
961 TTTCGAATGTCACGTTCCCATGCGACAAACCACCCGTGTG 71V-1658seq1
961 TTTCGAATGTCACGTTCCCATGCGACAAACCACCCGTGTG BFS-1703seq

                   1010       1020       1030       1040
1001 CTATTCACTGACGCCAGAACGAACACTCGACGTGCTCGAA 71V-1658seq1
1001 CTATTCACTGGCGCCAGAACGAACACTCGACGTGCTCGAA BFS-1703seq

                   1050       1060       1070       1080
1041 GAGAACGTCGACAATCCAAATTACGACACGCTGCTGGAGA 71V-1658seq1
1041 GAGAACGTCGACAATCCAAATTACGACACGCTGCTGGAGA BFS-1703seq

                   1090       1100       1110       1120
1081 ACGTCTTGAAATGTCCATCACGCCGGCCCAAACGAAGCAT 71V-1658seq1
1081 ACGTCTTGAAATGTCCATCACGCCGGCCCAAACGAAGCAT BFS-1703seq

                   1130       1140       1150       1160
1121 TACCGATGACTTCACACTGACCAGTCCCTACCTGGGGTTC 71V-1658seq1
1121 TACCGATGACTTCACACTGACCAGTCCCTACCTGGGGTTC BFS-1703seq

                   1170       1180       1190       1200
1161 TGCCCGTATTGCAGACACTCAACGCCGTGTTTCAGCCCAA 71V-1658seq1
1161 TGCCCGTATTGCAGACACTCAGCGCCGTGTTTCAGCCCAA BFS-1703seq

                   1210       1220       1230       1240
1201 TAAAAATTGAGAACGTGTGGGACGAATCTGATGATGGATC 71V-1658seq1
1201 TAAAAATTGAGAACGTGTGGGACGAATCTGATGATGGATC BFS-1703seq

                   1250       1260       1270       1280
1241 GATTAGAATCCAGGTCTCGGCACAATTCGGCTACAATCAG 71V-1658seq1
1241 GATTAGAATCCAGGTCTCGGCACAATTCGGCTACAATCAG BFS-1703seq

                   1290       1300       1310       1320
1281 GCAGGCACTGCGGATGTCACCAAATTCCGTTACATGTCTT 71V-1658seq1
1281 GCAGGCACTGCAGATGTCACCAAATTCCGTTACATGTCTT BFS-1703seq
```

```
                    1330        1340        1350        1360
1321 TCGACCACGACCATGACATCAAGGAAGACAGTATGGAGAA 71V-1658seq1
1321 TCGACCACGACCATGACATCAAGGAAGACAGTATGGATAA BFS-1703seq

                    1370        1380        1390        1400
1361 AATAGCTATCAGCACATCTGGACCCTGCCGTCGTCTTGGC 71V-1658seq1
1361 AATAGCTATTAGCACATCTGGACCCTGCCGTCGTCTTGGC BFS-1703seq

                    1410        1420        1430        1440
1401 CACAAAGGGTACTTCCTGTTAGCTCAATGTCCTCCAGGTG 71V-1658seq1
1401 CACAAAGGGTACTTCCTGTTAGCTCAATGTCCTCCAGGTG BFS-1703seq

                    1450        1460        1470        1480
1441 ACAGTGTAACCGTCAGTATCACGAGCGGAGCATCTGAGAA 71V-1658seq1
1441 ACAGTGTAACCGTCAGTATCACGAGCGGAGCATCTGAGAA BFS-1703seq

                    1490        1500        1510        1520
1481 TTCATGCACCGTGGAGAAAAAAGATCAGGAGGAAGTTTGTC 71V-1658seq1
1481 TTCATGCACCGTGGAGAAAAAAGATCAGGAGGAAGTTTGTC BFS-1703seq

                    1530        1540        1550        1560
1521 GGTAGAGAGGAGTACTTGTTCCCACCCGTCCATGGAAAGC 71V-1658seq1
1521 GGTAGAGAGGAGTACTTGTTCCCACCTGTCCATGGAAAGC BFS-1703seq

                    1570        1580        1590        1600
1561 TGGTAAAGTGCCACGTTTACGATCACTTGAAGGAGACGTC 71V-1658seq1
1561 TGGTAAAGTGCCACGTTTACGATCACTTGAAGGAGACGTC BFS-1703seq

                    1610        1620        1630        1640
1601 TGCCGGGTACATAACCATGCACAGGCCAGGCCCACACGCG 71V-1658seq1
1601 TGCCGGATACATAACCATGCACAGGCCAGGCCCACACGCG BFS-1703seq

                    1650        1660        1670        1680
1641 TATAAGTCCTATCTGGAGGAAGCGTCAGGCGAAGTGTACA 71V-1658seq1
1641 TATAAGTCCTATCTGGAGGAAGCGTCAGGCGAAGTGTACA BFS-1703seq

                    1690        1700        1710        1720
1681 TTAAACCACCTTCTGGCAAGAACGTCACCTACGAATGTAA 71V-1658seq1
1681 TTAAACCACCTTCTGGCAAGAACGTCACCTACGAATGTAA BFS-1703seq

                    1730        1740        1750        1760
1721 GTGTGGCGACTACAGCACAGGTATCGTGAGCACGCGAACG 71V-1658seq1
1721 GTGTGGCGACTACAGCACAGGTATTGTGAGCACGCGAACG BFS-1703seq
```

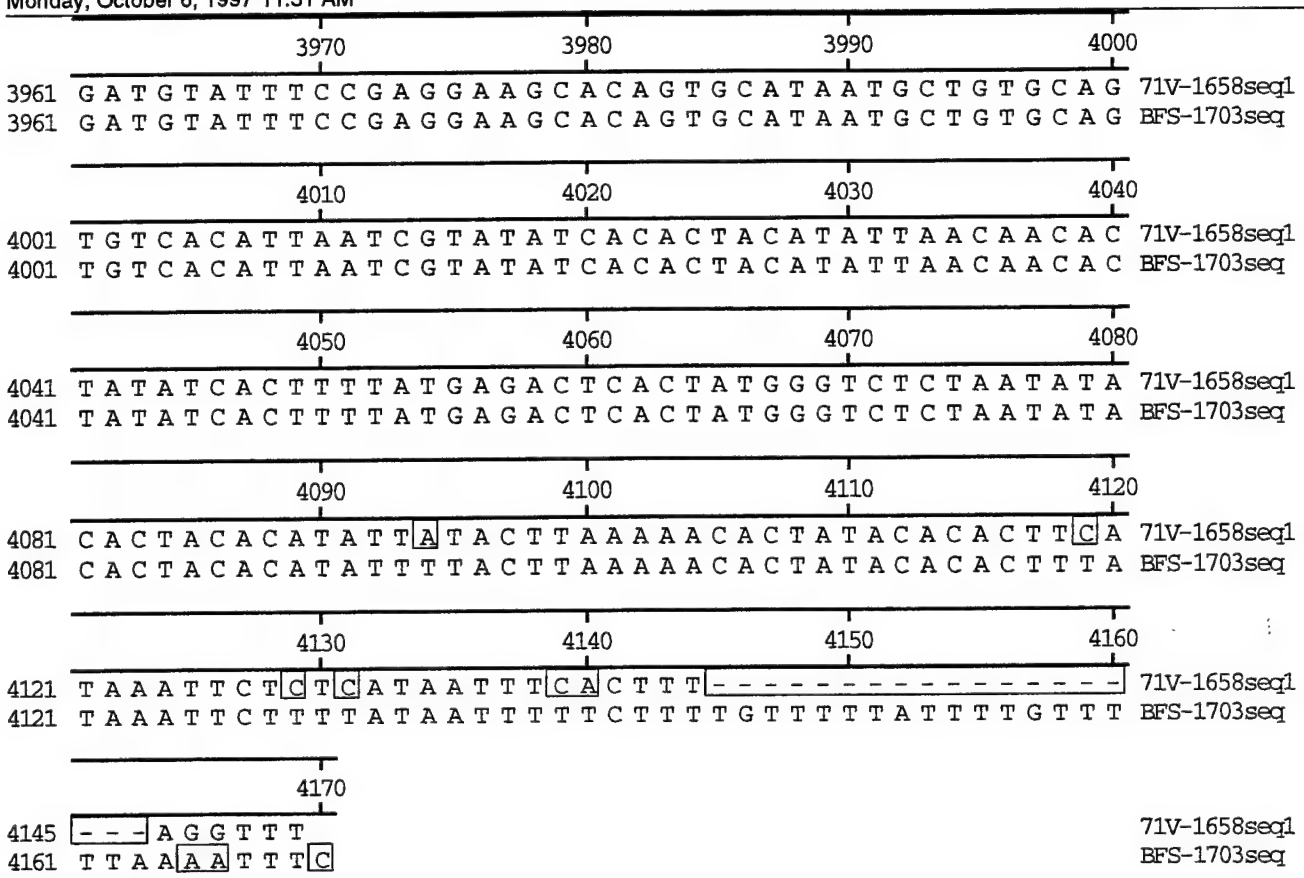

1770 1780 1790 1800
1761 A A G A T G A A C G G C T G C A C T A A A G C A A A A C A G T G C A T T G C C T 71V-1658seq1
1761 A A G A T G A A C G G C T G C A C T A A A G C A A A A C A G T G C A T T G C C T BFS-1703seq
1810 1820 1830 1840
1801 A C A A G A G C G A C C A A A C G A A A T G G G T C T T C A A C T C G C C G G A 71V-1658seq1
1801 A C A A G A G C G A C C A A A C G A A A T G G G T C T T C A A C T C G C C G G A BFS-1703seq
1850 1860 1870 1880
1841 T C T T A T T A G G C A C A C A G A C C A C T C A G T G C A A G G T A A A T T G 71V-1658seq1
1841 T C T T A T T A G G C A C A C A G A C C A C T C A G T G C A A G G T A A A C T G BFS-1703seq
1890 1900 1910 1920
1881 C A C A T T C C A T T C C G C T T G A C A C C G A C A G T C T G C C C G G T T C 71V-1658seq1
1881 C A C A T T C C A T T C C G C T T G A C A C C G A C A G T C T G C C C G G T T C BFS-1703seq
1930 1940 1950 1960
1921 C G T T A G C T C A C A C G C C T A C A G T C A C G A A G T G G T T C A A A G G 71V-1658seq1
1921 C G T T A G C T C A C A C G C C T A C A G T C A C G A A G T G G T T C A A A G G BFS-1703seq
1970 1980 1990 2000
1961 C A T C A C C C T C C A C C T G A C T G C A A T G C G A C C A A C A T T G C T G 71V-1658seq1
1961 C A T C A C C C T C C A C C T G A C T G C A A C G C G A C C A A C A T T G C T G BFS-1703seq
2010 2020 2030 2040
2001 A C A A C G A G A A A A T T G G G G C T G C G A G C A G A C G C A A C A G C A G 71V-1658seq1
2001 A C A A C G A G A A A A T T G G G G C T G C G A G C A G A C G C A A C A G C A G BFS-1703seq
2050 2060 2070 2080
2041 A A T G G A T T A C A G G G T C T A C A T C C A G G A A T T T T T C T G T G G G 71V-1658seq1
2041 A A T G G A T T A C A G G G A C T A C A T C C A G G A A T T T T T C T G T G G G BFS-1703seq
2090 2100 2110 2120
2081 G C G A G A A G G G C T G G A G T A C G T A T G G G G T A A C C A T G A A C C A 71V-1658seq1
2081 G C G A G A A G G G C T G G A G T A C G T A T G G G G C A A C C A T G A A C C A BFS-1703seq
2130 2140 2150 2160
2121 G T C A G A G T C T G G G C C C A G G A G T C G G C A C C A G G C G A C C C A C 71V-1658seq1
2121 G T C A G A G T C T G G G C C C A G G A G T C G G C A C C A G G C G A C C C A C BFS-1703seq
2170 2180 2190 2200
2161 A T G G A T G G C C G C A T G A G A T C A T C A T C C A C T A T T A T C A T C G 71V-1658seq1
2161 A T G G A T G G C C G C A T G A G A T C A T C A T C C A C T A T T A T C A T C G BFS-1703seq

```
                2210                2220                2230                2240
2201 GCATCCAGTCTACACTGTCATTGTGCTGTGTGGTGTCGCT 71V-1658seq1
2201 GCATCCAGTCTACACTGTCATTGTGCTGTGTGGTGTCGCT BFS-1703seq

                2250                2260                2270                2280
2241 CTTGCTATCCTGGTAGGCACTGCATC[A]TCAGCAGCTTG[C]A 71V-1658seq1
2241 CTTGCTATCCTGGTAGGCACTGCATCGTCAGCAGCTTGTA BFS-1703seq

                2290                2300                2310                2320
2281 TCGCCAAAGCAAGAAGAGACTGCCTGACGCCATACGCGCT 71V-1658seq1
2281 TCGCCAAAGCAAGAAGAGACTGCCTGACGCCATACGCGCT BFS-1703seq

                2330                2340                2350                2360
2321 TGCACCGAACGCAACGGTACCCACAGCATTAGCG[G]TTTTG 71V-1658seq1
2321 TGCACCGAACGCAACGGTACCCACAGCATTAGC[A]GTTTTG BFS-1703seq

                2370                2380                2390                2400
2361 TGCTG[C]ATTCGGCCAACCAACGCTGAAACATTTGGAGAAA 71V-1658seq1
2361 TGCTGTATTCGGCCAACCAACGCTGAAACATTTGGAGAAA BFS-1703seq

                2410                2420                2430                2440
2401 CTTTGAACCATCTGTGGTTTAACAACCAACCGTTTCTCTG 71V-1658seq1
2401 CTTTGAACCATCTGTGGTTTAACAACCAACCGTTTCTCTG BFS-1703seq

                2450                2460                2470                2480
2441 GGCACAGTTGTGCATTCCTCTGGCAGCGCTTGTTATTCTG 71V-1658seq1
2441 GGCACAGTTGTGCAT[C]CCTCTGGCAGCGCTTGTTATTCTG BFS-1703seq

                2490                2500                2510                2520
2481 TTCCGCTGCTTTTCATGCTGCATGCCTTTTTTATTGGTTG 71V-1658seq1
2481 TTCCGCTGCTTTTCATGCTGCATGCCTTTTTTATTGGTTG BFS-1703seq

                2530                2540                2550                2560
2521 CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAACATGC 71V-1658seq1
2521 CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAACATGC BFS-1703seq

                2570                2580                2590                2600
2561 GACCACTGTGCCAAATGTTCCGGGGATCCCGTATAAGGCG 71V-1658seq1
2561 GACCACTGTGCCAAATGTTCCGGGGATCCCGTATAAGGCG BFS-1703seq

                2610                2620                2630                2640
2601 TTGGTCGAACG[C]GCAGGTTACGCGCCACTTAA[C]CTGGAGA 71V-1658seq1
2601 TTGGTCGAACGTGCAGGTTACGCGCCACTTAATCTGGAGA BFS-1703seq
```


2650 2660 2670 2680
2641 TCACGGTCGTCTCATCGGAATTAACACCTTCAACTAACAA 71V-1658seq1
2641 TCACTGTCGTCTCATCGGAATTAACACCCTCAACTAACAA BFS-1703seq
2690 2700 2710 2720
2681 GGAGTACGTGACCTGCAAATTCCACACAGTCATTCCTTCA 71V-1658seq1
2681 GGAGTACGTGACCTGCAGATTCCACACAGTCATTCCTTCA BFS-1703seq
2730 2740 2750 2760
2721 CCACAAGTTAAATGCTGCGGGTCCCTCGAGTGCAAGGCAT 71V-1658seq1
2721 CCACAAGTTAAATGCTGCGGGTCCCTCGAGTGTAAGGCAT BFS-1703seq
2770 2780 2790 2800
2761 CCTCAAAGGCGGATTACACATGCCGCGTTTTTGGCGGTGT 71V-1658seq1
2761 CCTCAAAAGCGGATTACACATGCCGCGTTTTTGGCGGTGT BFS-1703seq
2810 2820 2830 2840
2801 GTACCCTTTCATGTGGGGAGGCGCACAATGCTTCTGTGAC 71V-1658seq1
2801 GTACCCTTTCATGTGGGGAGGCGCACAGTGCTTCTGTGAC BFS-1703seq
2850 2860 2870 2880
2841 AGTGAGAACACACAACTGAGTGAGGCGTACGTCGAGTTCG 71V-1658seq1
2841 AGTGAGAACACACAACTGAGTGAGGCATACGTCGAGTTCG BFS-1703seq
2890 2900 2910 2920
2881 CTCCAGACTGCACTATAGATCACGCAGTCGCACTAAAAGT 71V-1658seq1
2881 CTCCAGACTGCACTATTGATCATGCAGTCGCACTAAAAGT BFS-1703seq
2930 2940 2950 2960
2921 TCACACAGCTGCTCTGAAAGTCGGCCTGCGTATAGTATAC 71V-1658seq1
2921 TCACACAGCTGCTCTGAAAGTCGGCCTGCGTATAGTATAC BFS-1703seq
2970 2980 2990 3000
2961 GGCAACACCACCGCGCACCTGGATACGTTTGTCAATGGCG 71V-1658seq1
2961 GGCAATACCACCGCGCACCTGGATACGTTCGTCAATGGCG BFS-1703seq
3010 3020 3030 3040
3001 TCACGCCAGGTTCCTCACGGGACCTGAAGGTCATAGCAGG 71V-1658seq1
3001 TCACGCCAGGTTCCTCACGGGACCTGAAGGTCATAGCAGG BFS-1703seq
3050 3060 3070 3080
3041 GCCGATATCAGCCGCTTTTTCACCCTTTGACCATAAGGTC 71V-1658seq1
3041 GCCGATATCAGCCGCTTTTTCACCCTTTGACCATAAGGTC BFS-1703seq

```
                    3090                3100                3110                3120
3081 GTCATCAGAAAGGGGCTTGTTTACAACTACGACTTCCCTG 71V-1658seq1
3081 GTCATCAGAAAGGGGCTTGTTTACAACTACGACTTCCCTG BFS-1703seq

                    3130                3140                3150                3160
3121 AGTATGGAGCTATGAAACCAGGAGCGTTCGGCGATATTCA 71V-1658seq1
3121 A[A]TATGGAGCTATGAAACCAGGAGCGTTCGGCGATATTCA BFS-1703seq

                    3170                3180                3190                3200
3161 AGCATCCTCGCTTGATGCTACAGACATAGTAGCCCGCACT 71V-1658seq1
3161 AGCATCCTCGCTTGATGC[C]ACAGACATAGTAGCCCGCACT BFS-1703seq

                    3210                3220                3230                3240
3201 GACATACGGCTGCTGAAGCCTTCTGTCAAGAACATCCACG 71V-1658seq1
3201 GACATACGGCTGCTGAAGCCTTCTGTCAAGAACATCCACG BFS-1703seq

                    3250                3260                3270                3280
3241 TCCCCTACACCCAAGCAGTATCAGGGTATGAAATGTGGAA 71V-1658seq1
3241 TCCCCTACACCCAAGCAGTATCAGGGTATGAAATGTGGAA BFS-1703seq

                    3290                3300                3310                3320
3281 GAACAACTCAGGACGACCCCTGCAAGAAACAGCACCATTT 71V-1658seq1
3281 GAACAACTCAGGACGACCCCTGCAAGAAACAGCACCATTT BFS-1703seq

                    3330                3340                3350                3360
3321 GGATGTAAAATTGAAGTGGAGCCTCTGCGAGCGTCTAACT 71V-1658seq1
3321 GGATGTAAAATTGAAGTGGAGCCTCTGCGAGCGTCTAACT BFS-1703seq

                    3370                3380                3390                3400
3361 GTGCTTACGGGCACATCCCTATCTCGATTGACATCCCTGA 71V-1658seq1
3361 GTGCTTACGGGCACATCCCTATCTCGATTGACATCCCTGA BFS-1703seq

                    3410                3420                3430                3440
3401 TGCAGCTTTTGTGAGATCATCAGAATCACCAACAATTTTA 71V-1658seq1
3401 [C]GCAGCTTTTGTGAGATCATCAGAATCACCAACAATTTTA BFS-1703seq

                    3450                3460                3470                3480
3441 GAAGTTAGCTGCACAGTAGCAGACTGCATTTATTCTGCAG 71V-1658seq1
3441 GAAGTTAGCTGCACAGTAGCAGACTGCATTTATTCTGCAG BFS-1703seq

                    3490                3500                3510                3520
3481 ACTTTGGTGGTTCTCTAACATTACAGTACAAAGCTGACAG 71V-1658seq1
3481 ACTTTGGTGGTTC[G]CTAACATTACAGTACAAAGCTGACAG BFS-1703seq
```

```
                    3530                3540                3550                3560
3521 G G A G G G A C A T T G T C C A G T T C A C T C C C A C T C C A C G A C A G C T 71V-1658seq1
3521 G G A G G G A C A T T G T C C A G T T C A C T C C C A C T C C A C A A C A G C T BFS-1703seq

                    3570                3580                3590                3600
3561 G T T T T G A A G G A A G C G A C C A C A C A T G T G A C T G C C G T A G G C A 71V-1658seq1
3561 G T T T T G A A G G A A G C G A C C A C A C A T G T G A C T G C C G T A G G C A BFS-1703seq

                    3610                3620                3630                3640
3601 G C A T A A C A C T A C A T T T T A G C A C A T C G A G C C C A C A A G C A A A 71V-1658seq1
3601 G C A T A A C A C T A C A T T T T A G C A C A T C G A G C C C A C A A G C A A A BFS-1703seq

                    3650                3660                3670                3680
3641 T T T T A T A G T T T C G C T A T G C G G C A A G A A G T C C A C C T G C A A T 71V-1658seq1
3641 T T T T A T A G T T T C G C T A T G C G G C A A G A A G A C C A C C T G C A A T BFS-1703seq

                    3690                3700                3710                3720
3681 G C T G A A T G T A A A C C A C C G G C C G A C C A C A T A A T T G G A G A A C 71V-1658seq1
3681 G C T G A A T G T A A A C C A C C G G C C G A C C A C A T A A T T G G A G A A C BFS-1703seq

                    3730                3740                3750                3760
3721 C A C A T A A A G T C G A C C A A G A A T T C C A G G C G G C A G T T T C C A A 71V-1658seq1
3721 C A C A T A A A G T C G A C C A A G A A T T C C A G G C G G C A G T T T C C A A BFS-1703seq

                    3770                3780                3790                3800
3761 A A C A T C T T G G A A C T G G C T G C T T G C A C T G T T T G G G G G A G C A 71V-1658seq1
3761 A A C A T C T T G G A A T T G G C T G C T T G C A C T G T T T G G G G G A G C A BFS-1703seq

                    3810                3820                3830                3840
3801 T C A T C C C T C A T T G T T G T A G G A C T T A T A G T G T T G G T C T G C A 71V-1658seq1
3801 T C A T C C C T C A T T G T T G T A G G A C T T A T A G T G T T G G T C T G C A BFS-1703seq

                    3850                3860                3870                3880
3841 G C T C T A T G C T T A T A A A C A C A C G T A G A T G A C T G A G C G C G G A 71V-1658seq1
3841 G C T C T A T G C T T A T A A A C A C A C G T A G A T G A C T G A G C G C G G A BFS-1703seq

                    3890                3900                3910                3920
3881 C A C T G A C A T A G C G G T A A A A C T C G A T G T A C T T C C G A G G A A G 71V-1658seq1
3881 C A C T G A C A T A G C G G T A A A A C T C G A T G T A C T T C C G A G G A A G BFS-1703seq

                    3930                3940                3950                3960
3921 C G T G G T G C A T A A T G C C A C G C G C C G C T T G A C A C T A A A A C T C 71V-1658seq1
3921 C G T G G T G C A T A A C G C C A C G C G C C G C T T G A C A C T A A A A C T C BFS-1703seq
```



Decoration 'Decoration #1': Box residues that differ from the Consensus.

Appendix B Alignment of the amino acid sequences of the 26S region of WEE strain 71V-1658 and BFS-1703.

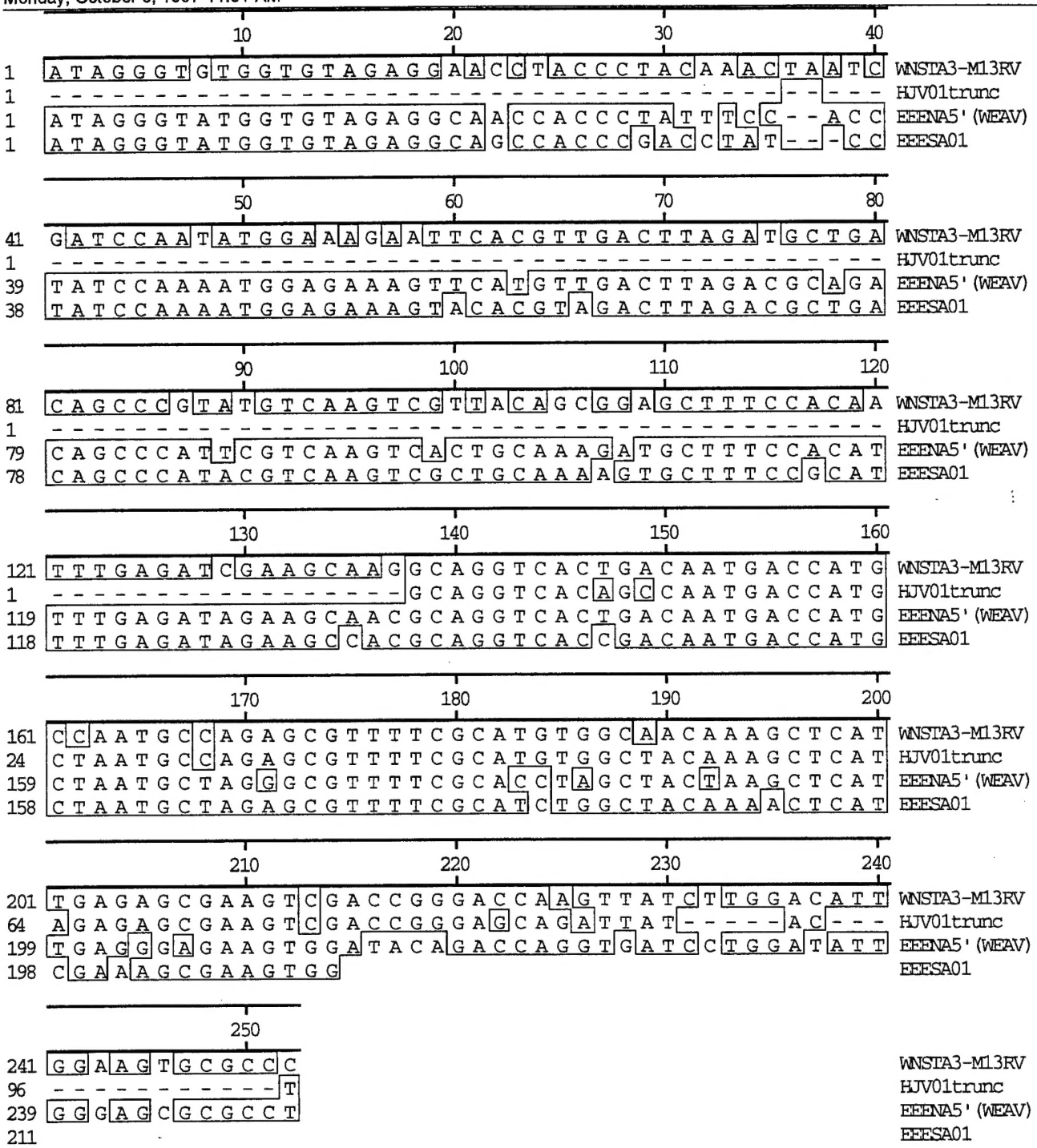
```
                    10        20        30        40
1   M F P Y P Q L N F P P V Y P T N P M A Y R D P N P P R R R W R P F R P P L A A Q  71V-1658prt
1   M F P Y P Q L N F P P V Y P T N P M A Y R D P N P P R C R W R P F R P P L A A Q  BFS-1703prt

                    50        60        70        80
41  I E D L R R S I V N L T F K Q R S P N P P P G P P P K K K S A P K P K P T Q P  71V-1658prt
41  I E D L R R S I A N L T F K Q R S P N P P P G P P P K K K S A P K P K P T Q P  BFS-1703prt

                    90        100       110       120
81  K K K K Q Q A K R T K R K P K P G K R Q R M C M K L E S D K T F P I M L N G Q V  71V-1658prt
81  K K K K Q Q A K K T K R K P K P G K R Q R M C M K L E S D K T F P I M L N G Q V  BFS-1703prt

                    130       140       150       160
121 N G Y A C V V G G R L M K P L H V E G K I D N E Q L A A V K L K K A S M Y D L E  71V-1658prt
121 N G Y A C V V G G R L M K P L H V E G K I D N E Q L A A V K L K K A S M Y D L E  BFS-1703prt

                    170       180       190       200
161 Y G D V P Q N M K S D T L Q Y T S D K P P G F Y N W H H G A V Q Y E N G R F T V  71V-1658prt
161 Y G D V P Q N M K S D T L Q Y T S D K P P G F Y N W H H G A V Q Y E N G R F T V  BFS-1703prt

                    210       220       230       240
201 P R G V G G K G D S G R P I L D N R G R V V A I V L G G A N E G T R T A L S V V  71V-1658prt
201 P R G V G G K G D S G R P I L D N R G R V V A I V L G G A N E G T R T A L S V V  BFS-1703prt

                    250       260       270       280
241 T W N Q K G V T I R D T P E G S E P W S L V T A L C V L S N V T F P C D K P P V  71V-1658prt
241 T W N Q K G V T I K D T P E G S E P W S L V T A L C V L S N V T F P C D K P P V  BFS-1703prt

                    290       300       310       320
281 C Y S L T P E R T L D V L E E N V D N P N Y D T L L E N V L K C P S R R P K R S  71V-1658prt
281 C Y S L A P E R T L D V L E E N V D N P N Y D T L L E N V L K C P S R R P K R S  BFS-1703prt

                    330       340       350       360
321 I T D D F T L T S P Y L G F C P Y C R H S T P C F S P I K I E N V W D E S D D G  71V-1658prt
321 I T D D F T L T S P Y L G F C P Y C R H S A P C F S P I K I E N V W D E S D D G  BFS-1703prt

                    370       380       390       400
361 S I R I Q V S A Q F G Y N Q A G T A D V T K F R Y M S F D H D H D I K E D S M E  71V-1658prt
361 S I R I Q V S A Q F G Y N Q A G T A D V T K F R Y M S F D H D H D I K E D S M D  BFS-1703prt

                    410       420       430       440
401 K I A I S T S G P C R R L G H K G Y F L L A Q C P P G D S V T V S I T S G A S E  71V-1658prt
401 K I A I S T S G P C R R L G H K G Y F L L A Q C P P G D S V T V S I T S G A S E  BFS-1703prt
```

```
                    450       460       470       480
441 NSCTVEKKIRRKFVGREEYLFPPVHGKLVKCHVYDHLKET 71V-1658prt
441 NSCTVEKKIRRKFVGREEYLFPPVHGKLVKCHVYDHLKET BFS-1703prt

                    490       500       510       520
481 SAGYITMHRPGPHAYKSYLEEASGEVYIKPPSGKNVTYEC 71V-1658prt
481 SAGYITMHRPGPHAYKSYLEEASGEVYIKPPSGKNVTYEC BFS-1703prt

                    530       540       550       560
521 KCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSP 71V-1658prt
521 KCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSP BFS-1703prt

                    570       580       590       600
561 DLIRHTDHSVQGKLHIPFRLTPTVCPVPLAHTPTVTKWFK 71V-1658prt
561 DLIRHTDHSVQGKLHIPFRLTPTVCPVPLAHTPTVTKWFK BFS-1703prt

                    610       620       630       640
601 GITLHLTA[M]RPTLLTTRKLGLRADATAEWITG[S]TSRNFSV 71V-1658prt
601 GITLHLTATRPTLLTTRKLGLRADATAEWITG[T]TSRNFSV BFS-1703prt

                    650       660       670       680
641 GREGLEYVWGNHEPVRVWAQESAPGDPHGWPHEIIIHYYH 71V-1658prt
641 GREGLEYVWGNHEPVRVWAQESAPGDPHGWPHEIIIHYYH BFS-1703prt

                    690       700       710       720
681 RHPVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYA 71V-1658prt
681 RHPVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYA BFS-1703prt

                    730       740       750       760
721 LAPNATVPTALAVLCCIRPTNAETFGETLNHLWFNNQPFL 71V-1658prt
721 LAPNATVPTALAVLCCIRPTNAETFGETLNHLWFNNQPFL BFS-1703prt

                    770       780       790       800
761 WAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAFEH 71V-1658prt
761 WAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAFEH BFS-1703prt

                    810       820       830       840
801 ATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTN 71V-1658prt
801 ATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTN BFS-1703prt

                    850       860       870       880
841 KEYVTC[K]FHTVIPSPQVKCCGSLECKASSKADYTCRVFGG 71V-1658prt
841 KEYVTC[R]FHTVIPSPQVKCCGSLECKASSKADYTCRVFGG BFS-1703prt
```

```
                  890       900       910       920
881  V Y P F M W G G A Q C F C D S E N T Q L S E A Y V E F A P D C T I D H A V A L K 71V-1658prt
881  V Y P F M W G G A Q C F C D S E N T Q L S E A Y V E F A P D C T I D H A V A L K BFS-1703prt

                  930       940       950       960
921  V H T A A L K V G L R I V Y G N T T A H L D T F V N G V T P G S S R D L K V I A 71V-1658prt
921  V H T A A L K V G L R I V Y G N T T A H L D T F V N G V T P G S S R D L K V I A BFS-1703prt

                  970       980       990      1000
961  G P I S A A F S P F D H K V V I R K G L V Y N Y D F P E Y G A M K P G A F G D I 71V-1658prt
961  G P I S A A F S P F D H K V V I R K G L V Y N Y D F P E Y G A M K P G A F G D I BFS-1703prt

                 1010      1020      1030      1040
1001 Q A S S L D A T D I V A R T D I R L L K P S V K N I H V P Y T Q A V S G Y E M W 71V-1658prt
1001 Q A S S L D A T D I V A R T D I R L L K P S V K N I H V P Y T Q A V S G Y E M W BFS-1703prt

                 1050      1060      1070      1080
1041 K N N S G R P L Q E T A P F G C K I E V E P L R A S N C A Y G H I P I S I D I P 71V-1658prt
1041 K N N S G R P L Q E T A P F G C K I E V E P L R A S N C A Y G H I P I S I D I P BFS-1703prt

                 1090      1100      1110      1120
1081 D A A F V R S S E S P T I L E V S C T V A D C I Y S A D F G G S L T L Q Y K A D 71V-1658prt
1081 D A A F V R S S E S P T I L E V S C T V A D C I Y S A D F G G S L T L Q Y K A D BFS-1703prt

                 1130      1140      1150      1160
1121 R E G H C P V H S H S T T A V L K E A T T H V T A V G S I T L H F S T S S P Q A 71V-1658prt
1121 R E G H C P V H S H S T T A V L K E A T T H V T A V G S I T L H F S T S S P Q A BFS-1703prt

                 1170      1180      1190      1200
1161 N F I V S L C G K K S T C N A E C K P P A D H I I G E P H K V D Q E F Q A A V S 71V-1658prt
1161 N F I V S L C G K K T T C N A E C K P P A D H I I G E P H K V D Q E F Q A A V S BFS-1703prt

                 1210      1220      1230
1201 K T S W N W L L A L F G G A S S L I V V G L I V L V C S S M L I N T R R .  71V-1658prt
1201 K T S W N W L L A L F G G A S S L I V V G L I V L V C S S M L I N T R R .  BFS-1703prt
```

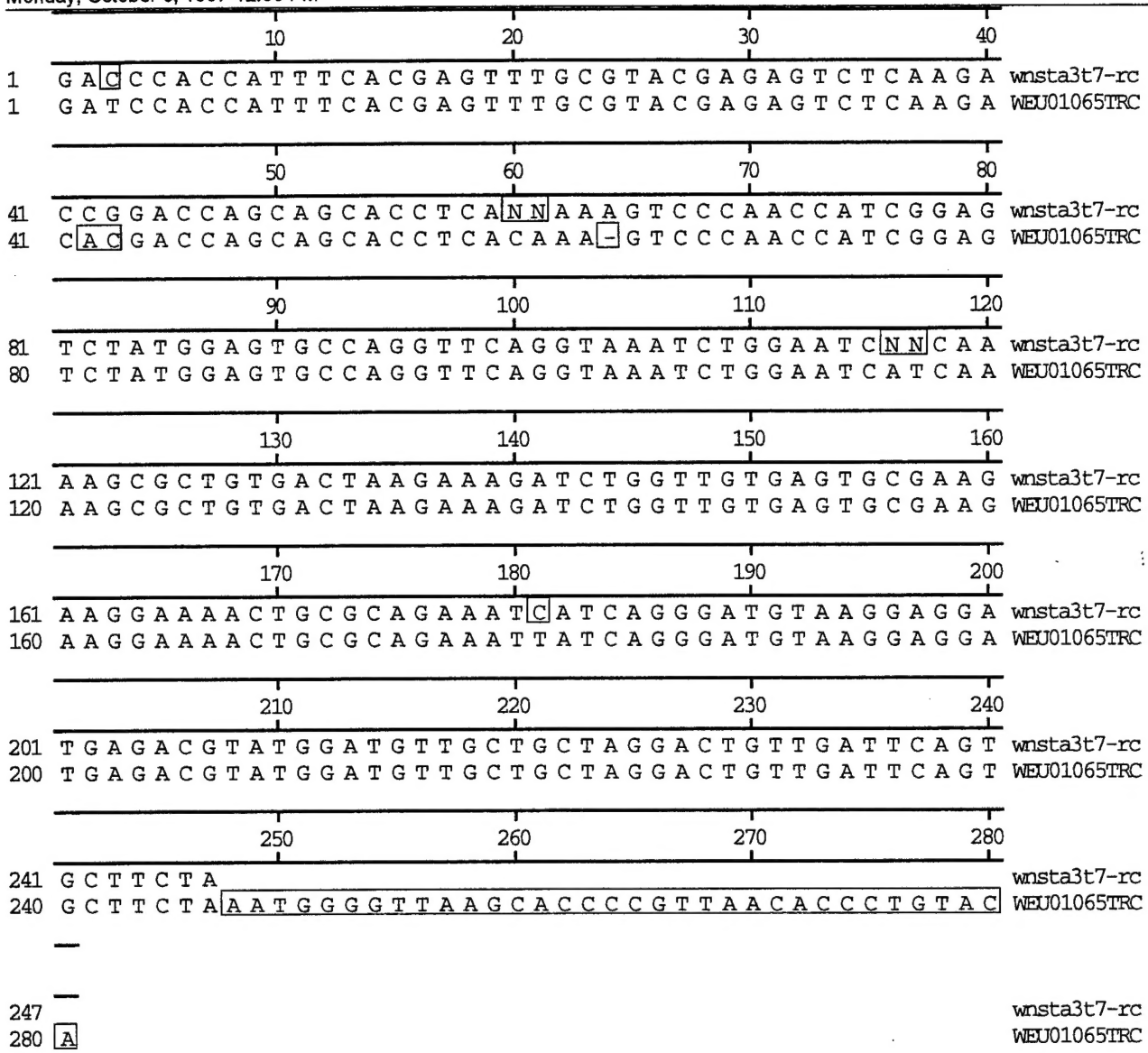
Decoration 'Decoration #1': Box residues that differ from the Consensus.

Appendix C a) Sequence of the 5' end of WNSTA-3 (WEE 71V-1658) aligned with the 5' ends of HJ (HJV01trunc), EEE NA [EEENA5'(WEAV)] and EEE SA (EEESA01); and b) Sequence of the 3' end of WNSTA-3 (WEE 71V-1658) aligned with the 5' end of WEE strain BFS 1703 (Weaver, *et al*, 1993) as described in text.



Decoration 'Decoration #1': Box residues that match the Consensus exactly.

Decoration 'Decoration #2': Box residues that match the Consensus exactly.



Decoration 'Decoration #1': Box residues that differ from the Consensus.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for who the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in Section 8.)

   Defence Research Establishment Suffield
   Box 4000, Medicine Hat, AB T1A 8K6

2. SECURITY CLASSIFICATION
   (overall security classification of the document, including special warning terms if applicable)

   UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title).

   Molecular Cloning and Sequencing of the 26S Region of Western Equine Encephalitis Virus Strain 71V-1658

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

   Schmaltz, F.L.; Nagata, L..P. and Rayner, G.

5. DATE OF PUBLICATION (month and year of publication of document)

   January 1997

6a. NO. OF PAGES (total containing information, include Annexes, Appendices, etc)

   66

6b. NO. OF REFS (total cited in document)

   33

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

   Suffield Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

   Defence Research Establishment Suffield

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

   Thrust 6DB Vaccine Development

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)
   SR 633

10b. OTHER DOCUMENT NOs. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

   (X) Unlimited distribution
   ( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
   ( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
   ( ) Distribution limited to government departments and agencies; further distribution only as approved
   ( ) Distribution limited to defence departments; further distribution only as approved
   ( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally corresponded to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected).

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C) or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

A 3100 bp cDNA clone (pcDW-12) of the 26S region of western equine encephalitis (WEE) virus strain 71V-1658 was identified by dot blot hybridization from a cDNA library. The missing 5' end was PCR-amplified engineered into pcDW-12 to obtained a full length clone of the 26S region. The resulting construct (XH-7) was restriction mapped and completely sequenced on both strands. Only eleven of the sixty-three nucleotide differences resulted in amino acid changes when the sequence was compared to WEE strain BFS-1703 (Hahn *et al*, 1988). In addition, the high degree of conservation of the structural proteins was maintained when compared to the N-terminal sequence of the E1 and E2 proteins of the McMillan strain of WEE (Bell *et al*, 1983). The conserved nature of the structural proteins would indicate that one strain of WEE should be able to cross-protect against all WEE strains. A 2.2 kb fragment at the 5' end of the genome was also cloned and sequenced, and demonstrated high homology to the available sequence for eastern equine encephalitis (EEE) and Highlands J (HJ) viruses, adding further evidence that the entire 5' nonstructural region of WEE was derived from EEE. In summary, the genetic cloning and sequencing of the WEE 26S region marks the critical first step in the generation of a subunit vaccine to WEE.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifies, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

western equine encephalitis virus

subunit vaccines

alphavirus